## Air Toxics Hot Spots Program
## Risk Assessment Guidelines

## Part II

# Technical Support Document for Cancer Potency Factors

**June 2008** **Public Review draft**


Secretary for Environmental Protection
California Environmental Protection Agency
Linda S. Adams

Director
Office of Environmental Health Hazard Assessment
Joan E. Denton, Ph.D.

# Technical Support Document for Cancer Potency Factors:

**Methodologies for derivation, listing of available values, and adjustments to allow for early life stage exposures.**

**2008**

**California Environmental Protection Agency**

**Office of Environmental Health Hazard Assessment**

**Air Toxicology and Epidemiology Branch**

Prepared by:

John D. Budroe, Ph.D.

James F. Collins, Ph.D.

Melanie A. Marty, Ph.D.

Andrew G. Salmon, M.A., D. Phil.

Joseph P. Brown, Ph.D.

Air Toxicology and Epidemiology Branch

and

Martha S. Sandy, Ph.D., M.P.H.

Claire D. Sherman, Ph.D.

Rajpal S. Tomar, Ph.D.

Lauren Zeise, Ph.D.

Reproductive and Cancer Hazard Assessment Branch,

Office of Environmental Health Hazard Assessment

Reviewed By

George V. Alexeeff, Ph.D.

Melanie A. Marty, Ph.D.

Lauren Zeise, Ph.D.

# EXECUTIVE SUMMARY

The Air Toxics "Hot Spots" Information and Assessment Act (AB 2588, Connelly) was enacted in September 1987. Under this Act, stationary sources of air pollution are required to report the types and quantities of certain substances their facilities routinely release into the air. The goals of the Air Toxics "Hot Spots" Act are to collect emission data, identify facilities having localized impacts, ascertain health risks posed by those facilities, notify nearby residents of significant risks and reduce emissions from significant sources.

The Technical Support Document for Cancer Potency Factors (TSD) contains cancer unit risks and potency factors for 107 of the 201 carcinogenic substances or groups of substances for which emissions must be quantified in the Air Toxics Hot Spots program. These unit risks are used in the cancer risk assessment of facility emissions.

The purpose of this document is to provide a summary of the data supporting the carcinogenic potential of individual substances or groups of substances, to provide the calculation procedures used to derive the estimated unit risk and cancer potency factors and describe procedures to consider the increased susceptibility of infants and children compared to adults to carcinogens.

The procedures used to consider the increased susceptibility of infants and children compared to adults to carcinogens include the use of age-specific weighting factors in calculating cancer risks from exposures of infants, children and adolescents, to reflect their anticipated special sensitivity to carcinogens. These procedures will not be used to impose any overall revisions of the existing cancer potencies and unit risks.

Additionally, this revision of the cancer TSD includes an update of the calculation procedures used to derive the estimated unit risk and cancer potency factors, and a cancer unit risk and data summary for ethyl benzene.

This document is one part of the Air Toxics Hot Spots Program Risk Assessment Guidelines. The other documents originally included in the Guidelines are Part I: Technical Support Document for the Determination of Acute Toxicity Reference Exposure Levels for Airborne Toxicants; Part III: Technical Support Document for Determination of Noncancer Chronic Reference Exposure Levels; Part IV: Technical Support Document for Exposure Assessment and Stochastic Analysis; Part V: Air Toxic Hot Spots Program Risk Assessment Guidelines. As a part of the same revision process which led to production of this revised TSD on cancer potencies, the original TSDs for Acute and Chronic Reference Exposure Levels are in the process of being replaced with a new unified TSD for Acute, 8-hour and Chronic Reference Exposure Levels.

## PREFACE

The Air Toxics "Hot Spots" Information and Assessment Act (AB 2588, Connelly) was enacted in September 1987. Under this Act, stationary sources are required to report the types and quantities of certain substances their facilities routinely release into the air. The goals of the Air Toxics "Hot Spots" Act are to collect emission data, identify facilities having localized impacts, ascertain health risks posed by those facilities, notify nearby residents of significant risks and reduce emissions from significant sources.

The Technical Support Document for Cancer Potency Factors (TSD) contains cancer unit risks and potency factors for 107 of the 201 carcinogenic substances or groups of substances for which emissions must be quantified in the Air Toxics Hot Spots program. These unit risks are used in risk assessment of facility emissions. The purpose of this document is to provide a summary of the data supporting the carcinogenic potential of individual substances or groups of substances, to provide the calculation procedure used to derive the estimated unit risk and cancer potency factors and describe procedures to consider early-life susceptibility to carcinogens.

In this document, OEHHA is responding to the requirements of the 1999 Children's Environmental Health Protection Act, (SB25, Escutia) by revising the procedures for derivation and application of cancer potency factors to take account of general or chemical-specific information which suggests that children may be especially susceptible to certain carcinogens (OEHHA, 2001a). The revised cancer potency derivation procedures described will not be used to impose any overall revisions of the existing cancer potencies, although they do reflect updated methods of derivation. However, individual cancer potency values will also be reviewed as part of the ongoing re-evaluation of health values mandated by SB 25, and revised values will be listed in updated versions of the appendices to this document as necessary. The revisions also include the use of weighting factors in calculating cancer risks from exposures of infants, children and adolescents, to reflect their anticipated special sensitivity to carcinogens. Similar legal mandates to update risk assessment methodology and cancer potencies apply to the OEHHA program for development of Public Health Goals (PHGs) for chemicals in drinking water, and Proposition 65 No Significant Risk Levels (NSRLs); the latter may also be revised to reflect concerns for children's health. Revising these numbers will require the originating program to reconsider the value in an open public process. For example, OEHHA would need to release any revised potency factors for public comment and review by the Scientific Review Panel on Toxic Air Contaminants (SRP) prior to adoption under the TAC program. In addition, the procedures for outside parties to request reevaluation of cancer potency values by the programs which originated those values are listed in Appendix G.

This document lists adopted Cal/EPA values which were included in the previous version of the TSD for Cancer Potency Factors (OEHHA, 2005). Cal/EPA values were developed under the Toxic Air Contaminant (TAC) program, the PHG program, the Proposition 65 program, or in some cases specifically for the Air Toxics Hot Spots program. All the Cal/EPA values are submitted for public comments and external peer review prior to adoption by the program of origin.

OEHHA uses a toxic equivalency factor procedure for dioxin-like compounds, including polychlorinated dibenzo-*p*-dioxins, dibenzofurans and polychlorinated biphenyls (PCBs). The Toxicity Equivalency Factor scheme ($TEF_{WHO-97}$) developed by the World Health Organization/European Center for Environmental Health (WHO-ECEH) is used for determining cancer unit risk and potency values for these chemicals where individual congener emissions are available (Appendix C).

Some U.S. EPA IRIS cancer unit risk values were adopted under the previous versions of these guidelines, and these values will continue to be used unless and until revised by Cal/EPA. U.S. EPA has recently revised its cancer risk assessment guidelines (U.S. EPA, 2005a). Some of the recommended changes in methodology could result in slightly different potency values compared to those calculated by the previous methodology, although in practice a number of the recommendations (for example, the use of ¾ power of the body weight ratio rather than ⅔ power for interspecies scaling) have been available in draft versions of the revised policy for some time and appear in many more recent assessments. U.S. EPA has stated that cancer potency values listed in IRIS will not be revisited solely for the purpose of incorporating changes in cancer potency value calculation methods contained in the revised cancer risk assessment guidelines. U.S. EPA has also issued supplementary guidelines on assessing cancer risk from early-life exposure (U.S. EPA, 2005b).

This document is one part of the Air Toxics Hot Spots Program Risk Assessment Guidelines. The other documents originally included in the Guidelines are Part I: Technical Support Document for the Determination of Acute Toxicity Reference Exposure Levels for Airborne Toxicants; Part III: Technical Support Document for Determination of Noncancer Chronic Reference Exposure Levels; Part IV: Technical Support Document for Exposure Assessment and Stochastic Analysis; Part V: Air Toxic Hot Spots Program Risk Assessment Guidelines. As a part of the same revision process which led to production of this revised TSD on cancer potencies, the original TSDs for Acute and Chronic Reference Exposure Levels are in the process of being replaced with a new unified TSD for Acute, 8-hour and Chronic Reference Exposure Levels.

**TABLE OF CONTENTS**

## APPENDICES

**Appendix A.** A lookup table containing unit risk and cancer potency values.

**Appendix B.** Chemical-specific summaries of the information used to derive unit risk and cancer potency values.

**Appendix C.** A description of the use of toxicity equivalency factors for determining unit risk and cancer potency factors for polychlorinated dibenzo-*p*-dioxins, dibenzofurans and dioxin-like polychlorinated biphenyls.

**Appendix D.** A listing of Toxic Air Contaminants identified by the California Air Resources Board.

**Appendix E.** Descriptions of the International Agency for Research on Cancer (IARC) and U.S. Environmental Protection Agency (U.S. EPA) carcinogen classifications.

**Appendix F.** An asbestos quantity conversion factor for calculating asbestos concentrations expressed as 100 fibers/$m^3$ from asbestos concentrations expressed as µg/$m^3$.

**Appendix G.** Procedures for revisiting or delisting cancer potency factors by the program of origin.

**Appendix H.** Exposure routes and studies used to derive cancer unit risks and slope factors.

**Appendix I.** "Assessing susceptibility from early-life exposure to carcinogens": Barton *et al*., 2005 (from *Environmental Health Perspectives)*.

**Appendix J.** "In Utero and Early Life Susceptibility to Carcinogens: The Derivation of Age-at-Exposure Sensitivity Measures" – conducted by OEHHA's Reproductive and Cancer Hazard Assessment Branch.

**Appendix K.** Additions and corrections from prior document versions.

## INTRODUCTION

The Technical Support Document (TSD) for Describing Available Cancer Potency Factors provides technical information support for the Air Toxics Hot Spots Program Risk Assessment Guidelines. The TSD consists of 12 sections:

1. The TSD introduction.

2. A description of the methodologies used to derive the unit risk and cancer potency values listed in the lookup table.

3. A lookup table containing unit risk and cancer potency values. (Appendix A)

4. Chemical-specific summaries of the information used to derive unit risk and cancer potency values. (Appendix B).

5. A description of the use of toxicity equivalency factors for determining unit risk and cancer potency factors for polychlorinated dibenzo-*p*-dioxins, dibenzofurans and dioxin-like polychlorinated biphenyls (Appendix C).

6. A listing of Toxic Air Contaminants identified by the California Air Resources Board (Appendix D).

7. Descriptions of the International Agency for Research on Cancer (IARC) and U.S. Environmental Protection Agency (U.S. EPA) carcinogen classifications (Appendix E).

8. An asbestos quantity conversion factor for calculating asbestos concentrations expressed as 100 fibers/$m^3$ from asbestos concentrations expressed as µg/$m^3$ (Appendix F).

9. Procedures for revisiting or delisting cancer potency factors by the program of origin (Appendix G).

10. Exposure routes and studies used to derive cancer unit risks and slope factors (Appendix H).

11. "Assessing susceptibility from early-life exposure to carcinogens": Barton *et al*., 2005 (from *Environmental Health Perspectives)* (Appendix I).

12. "In Utero and Early Life Susceptibility to Carcinogens: The Derivation of Age-at-Exposure Sensitivity Measures" – conducted by OEHHA's Reproductive and Cancer Hazard Assessment Branch (Appendix J)

## SELECTION OF CANCER POTENCY VALUES

Although the Office of Environmental Health Hazard Assessment (OEHHA) has developed a number of cancer potencies for use in the Toxic Air Contaminants and Air Toxics Hot Spots programs, this document also describes and adopts other numbers which were originally developed for other California Environmental Protection Agency (Cal/EPA) programs, or by the U.S. EPA. These were reviewed for accuracy, reliance on up-to-date data and methodology, and applicability in the context of the Air Toxics Hot Spots program; values found appropriate were adopted here rather than devoting the resources necessary for a full *de novo* assessment. In some instances, several governmental organizations have calculated cancer potency values for the same chemical. The methodology used to select sources for unit risk and cancer potency values included in the Technical Support Document (TSD) for Describing Available Cancer Potency Factors have been described in past versions of this document and will not be described here.

All the cancer potency value sources used generally follow the recommendations of the National Research Council on cancer risk assessment (NRC, 1983, 1994). All Cal/EPA program documents undergo a process of public comment and scientific peer review prior to adoption, although the procedures used vary according to the program. The publication procedure for Toxic Air Contaminant documents includes a public comment period and review by the Scientific Review Panel on Toxic Air Contaminants (SRP) before identification of a Toxic Air Contaminant by the Air Resources Board of the California Environmental Protection Agency (Cal/EPA). Furthermore, a petition procedure is available to initiate TAC document review and revision if appropriate because of new toxicity data. Documents developed for the Air Toxics Hot Spots program similarly undergo public comment and peer review by the SRP before adoption by the Director of OEHHA. The standard Proposition 65 document adoption procedure includes a public comment and external peer review by the Proposition 65 Carcinogen Identification Committee. The expedited Proposition 65 document adoption procedure included a public comment period. Risk assessments prepared for development of Public Health Goals (PHGs) for chemicals in drinking water are subject to two public comment periods before the final versions and responses to comments are published on the OEHHA Web site. PHG documents may also receive external peer review. Documents from U.S. EPA's Integrated Risk Information System (IRIS) receive external peer review and are posted on the Internet for public viewing during the external peer review period, and any public comments submitted are considered by the originating office. Additionally, public comment may be solicited during the document posting period. Future preference for use of developed cancer potency factors/unit risks in this document will be done on a case by case basis, with preference given to those assessments most relevant to inhalation exposures of the California population, to the most recent derivations using the latest data sets and scientific methodology, and to those having undergone the most open and extensive peer review process.

## CANCER RISK ASSESSMENT METHODOLOGIES

This section describes in general the methodologies used to derive the cancer unit risk and potency factors listed in this document. As noted in the Preface to this document, no new cancer unit risks or potency factors were developed for this document; rather, all of the values contained here were previously developed in documents by Cal/EPA or U.S. EPA. Following the recommendations of the National Academy of Sciences (NRC, 1983), Cal/EPA and U.S. EPA have both used formalized cancer risk assessment guidelines, the original versions of which (California Department of Health Services, 1985; U.S. EPA, 1986) were published some time ago. Both these guidelines followed similar methodologies.

In the twenty years since these original guidelines were published there have been a number of advances in the methodology of cancer risk assessment, as well as very considerable advances in the quantity of data available not only from animal carcinogenesis bioassays and epidemiological studies, but also from mechanistic studies of carcinogenesis and related phenomena. Some of these advances have been incorporated into newer risk assessments by both agencies on a more or less *ad hoc* basis, but there has also been an ongoing effort to provide updated risk assessment guidance documents. In 1995, U.S. EPA released for public comment the "Proposed and Interim Guidelines for Carcinogen Risk Assessment", which was the first of several drafts released for public comment. Many risk assessments appearing since then have used elements of the recommendations contained in that document, in spite of its draft status. A final version of the U.S. EPA's revised cancer risk assessment guidelines has now been released (U.S. EPA, 2005a). Although these new guidelines incorporate a number of substantial changes from their predecessors (U.S. EPA, 1986; 1995), U.S. EPA has stated that cancer potency values listed in IRIS will not be revisited solely for the purpose of incorporating changes in cancer potency value calculation methods.

Cal/EPA has not produced a revised cancer risk assessment guideline document to replace the original version (DHS, 1985), but rather has relied on incorporating new data and methodologies as these became available, and described the methods used on a case by case basis in the individual risk assessment documents where these went beyond the original guidance. However, this revision of the TSD for cancer potencies provides a convenient opportunity to summarize the current status of the methodology used by OEHHA for the air toxics programs, and also to highlight points of similarity to, and difference from, the recommendations of U.S. EPA (2005a).

In this document, OEHHA intends to follow the recommendations of the NRC (1994) in describing a set of clear and consistent principles for choosing and departing from default cancer risk assessment options. NRC identified a number of objectives that should be taken into account when considering principles for choosing and departing from default options: "protecting the public health, ensuring scientific validity, minimizing serious errors in estimating risks, maximizing incentives for research, creating an orderly and predictable process, and fostering openness and trustworthiness". The OEHHA cancer risk methodologies discussed in this document are intended to generally meet those objectives cited above.

## Hazard Identification

This section will describe: 1) how weight of evidence evaluations are used in hazard evaluation; 2) guidelines for inferring causality of effect; 3) the use of human and animal carcinogenicity data, as well as supporting evidence (e.g. genetic toxicity and mechanistic data); 4) examples of carcinogen identification schemes.

*Evaluation of Weight of Evidence*

In evaluating the range of evidence on the toxicity and carcinogenicity of a compound, mixture or other agent, a "weight-of-evidence" approach is generally used to describe the body of evidence on whether or not exposure to the agent causes a particular effect. Under this approach, the number and quality of toxicological and epidemiological studies, as well as the consistency of study results and other sources of data on biological plausibility, are considered. Diverse and sometimes conflicting data need to be evaluated with respect to possible explanations of differing results. Consideration of methodological issues in the review of the toxicological and epidemiological literature is important in evaluating associations between exposure to an agent and animal or human health effects. This aspect of the evaluation process has received particular emphasis with respect to epidemiological data, where concerns as to the significance and reliability of the data and the impacts of confounding and misclassification are pressing. Such concerns are also relevant to some extent in the interpretation of animal bioassay data and mechanistic studies. Although the test animals, laboratory environment and characterization of the test agent are usually much better controlled than the equivalent parameters in an epidemiological study, the statistical issues of group sizes are likely to be less favorable, and there are uncertainties associated with extrapolation of biological responses from test animal species to humans.

*Criteria for Causality*

There has been extensive discussion over the last two centuries on causal inference. This has been particularly with regard to epidemiological data, but is also relevant to interpretation of animal studies. Most epidemiologists utilize causal inference guidelines based on those proposed by Bradford Hill (1971). OEHHA has relied on these and on recommendations by IARC (2006), the Institute of Medicine (2004), the Surgeon General's Reports on Smoking (U.S. DHHS, 2004) and standard epidemiologic texts (*e.g.* Lilienfeld and Lilienfeld, 1980; Rothman and Greenland, 1998). The criteria for determination of causality used by OEHHA have been laid out in various risk assessment documents: the summary below is adapted from the Health Effects section of the document prepared to support the identification of environmental tobacco smoke (ETS) as a Toxic Air Contaminant (OEHHA, 2005).

1. *Strength of Association*. A strong association between a factor and a disease (historically considered to be a relative risk or odds ratio $\geq 2$; and statistically significant) makes alternative explanations for the disease less likely. Small magnitude associations (i.e. risk estimate $> 1$ but $\leq 2$) make alternative explanations (undetected biases or confounders) more likely. However, such small magnitude associations do not necessarily indicate lack of causality and are relatively common in environmental

epidemiology. For example, the widely-accepted associations between air pollution and cardiovascular/pulmonary mortality, and passive smoking and lung cancer are considered small magnitude associations (risk estimate >1 and < 2). It is important to avoid confusing small magnitude of association with statistical insignificance. From a public health perspective, such small magnitude associations for a common disease can mean large numbers of people affected by the health outcome when exposure is frequent and widespread.

2. *Consistency of Association.* If several investigations find an association between a factor and a disease across a range of populations, geographic locations, times, and under different circumstances, then the factor is more likely to be causal. Consistency argues against hypotheses that the association is caused by some other factor(s) that varies across studies. Unmeasured confounding is an unlikely explanation when the effect is observed consistently across a number of studies in different populations.

   Associations that are replicated in several studies of the same design or using different epidemiological approaches or considering different sources of exposure and in a number of geographical regions are more likely to represent a causal relationship than isolated observations from single studies (IARC, 2006). If there are inconsistent results among investigations, possible reasons are sought (such as adequacy of sample size or control group, methods used to assess exposure, range in levels of exposure), and results of studies judged to be rigorous are emphasized over those of studies judged to be methodologically less rigorous. For example, studies with the best exposure assessment are more informative for assessing the association between ETS and breast cancer than studies with limited exposure assessment, all else being equal.

3. *Temporality.* Temporality means that the factor associated with causing the disease occurs in time prior to development of the disease.

4. *Coherence and Biological Plausibility.* A causal interpretation cannot conflict with what is known about the biology of the disease. The availability of experimental data or mechanistic theories consistent with epidemiological observations strengthens conclusions of causation. For example, the presence of known carcinogens in tobacco smoke supports the concept that exposure to tobacco smoke could cause increased cancer risk. Similarly, if the mechanism of action for a toxicant is consistent with development of a specific disease, then coherence and biological plausibility can be invoked. For example, cigarette smoke causes atherosclerosis, and atherosclerosis is involved in heart disease; thus, there is coherence with the epidemiologic finding that smoking elevates risk of heart disease.

5. *Dose-Response.* A basic tenet of toxicology is that increasing exposure or dose generally increases the response to the toxicant. While dose-response curves vary in shape and are not necessarily always monotonic, an increased gradient of response with increased exposure makes it difficult to argue that the factor is not associated with the disease. To argue otherwise necessitates that an unknown factor varies consistently with the dose of the substance and the response under question. While increased risk with increasing

levels of exposure is considered to be a strong indication of causality, absence of a graded response does not exclude a causal relationship (IARC, 2006).

The dose-response curves for specific toxic effects may be non-monotonic. Under appropriate circumstances, where the dose response shows saturation, the effect of exposures could be nearly maximal, with any additional exposure having little or no effect. For example, a dose response is observed for various cardiovascular endpoints in the range of exposures characteristic of environmental tobacco smoke, but for some the magnitude of the response is little different between active and passive smoking.

6. *Specificity.* Specificity is generally interpreted to mean that a single cause is associated with a single effect. It may be useful for determining which microorganism is responsible for a particular disease, or associating a single carcinogenic chemical with a rare and characteristic tumor (e.g., liver angiosarcoma and vinyl chloride, or mesothelioma and asbestos). But it is not helpful when studying diseases that are multifactorial, or toxic substances that contain a number of individual constituents, each of which may have several effects and/or target sites. Thus, specificity is not particularly relevant to the evaluation of health effects of tobacco smoke.

7. *Experimental evidence.* While experiments are often conducted over a short period of time or under artificial conditions (compared to real-life exposures), experiments offer the opportunity to collect data under highly controlled conditions that allow strong causal conclusions to be drawn. Experimental data that are consistent with epidemiological results strongly support conclusions of causality. There are also "natural experiments" that can be studied with epidemiological methods, such as when exposure of a human population to a substance declines or ceases; if the effect attributed to that exposure decreases, then there is evidence of causality. One example of this is the drop in heart disease death and lung cancer risk after smoking cessation.

It should be noted that the causal criteria are guidelines for judging whether a causal association exists between a factor and a disease, rather than hard-and-fast rules. Lilienfeld and Lilienfeld (1980) note that "*In medicine and public health, it would appear reasonable to adopt a pragmatic concept of causality. A causal relationship would be recognized to exist whenever evidence indicates that the factors form part of the complex of circumstances that increases the probability of the occurrence of disease and that a diminution of one or more of these factors decreases the frequency of that disease. After all, the reason for determining the etiological factors of a disease is to apply this knowledge to prevent the disease.*"

<u>Data sources</u>

*Human studies: epidemiology, ecological studies and case reports*

The aim of a risk assessment for the California Air Toxics programs is to determine potential impact on human health. Ideally therefore, the hazard identification would rely on studies in humans to demonstrate the nature and extent of the hazard. However, apart from clinical trials of drugs, experimental studies of toxic effects in human subjects are rarely undertaken or justifiable. Pharmacokinetic studies using doses below the threshold for any toxic effect have been undertaken for various environmental and occupational agents, but are not usually regarded as appropriate for suspected carcinogens.

The human data on carcinogens available to the risk assessor therefore mostly consist of epidemiological studies of existing occupational or environmental exposures. It is easier to draw reliable inferences in situations where both the exposures and the population are substantial and well-defined, and accessible to direct measurement rather than recall. Thus, many important findings of carcinogenicity to humans are based on analysis of occupational exposures. Problems in interpretation of occupational epidemiological data include simultaneous exposure to several different known or suspected carcinogens, imprecise quantification of exposures and confounding exposures such as active or passive tobacco smoking. The historical database of occupational data has a bias in favor of healthy adult males. Thus, the hazard analysis of these studies may not accurately characterize effects on women, infants, children or the elderly, or on members of minority ethnic groups. Nevertheless, the analysis of occupational epidemiological studies has proved an important source for unequivocal identification of human carcinogens.

Epidemiological evidence may also be obtained where a substantial segment of a general population is exposed to the material of interest in air, drinking water or food sources. Rigorous cohort and case-control studies may sometimes be possible, in which exposed individuals are identified, their exposure and morbidity or mortality evaluated, and compared to less exposed but otherwise similar controls. More often at least the initial investigation is a cross-sectional study, where prevalence of exposures and outcomes is compared in relatively unexposed and exposed populations. Such studies are hypothesis-generating, but are important sources of information nevertheless, and can often also justify more costly and labor-intensive follow-up cohort and/or case-control studies.

The clinical medical literature contains many case reports where a particular health outcome is reported along with unusual exposures that might have contributed to its occurrence. These reports typically describe a single patient or a small group, and have no statistical significance. They are nevertheless useful as indications of possible associations that deserve follow-up using epidemiological methods, and as supporting evidence, addressing the plausibility of associations measured in larger studies.

*Animal studies*

Although the observation of human disease in an exposed population can provide definitive hazard identification, adequate data of this type are not always available. More often, risk estimates have to be based on studies in experimental animals, and extrapolation of these results to predict human toxicity. The animals used are mostly rodents, typically the common laboratory strains of rat and mouse.

Rats and mice have many similarities to humans. Physiology and biochemistry are similar for all mammals, especially at the fundamental levels of xenobiotic metabolism, DNA replication and DNA repair that are of concern in identifying carcinogens. However, there are also several important differences between rodents and humans. Rodents, with a short life span, have differences in cell growth regulation compared to longer-lived species such as the human. For instance, whereas laboratory investigations have suggested that mutations in two regulatory genes (*e.g.* H-ras and p-53) are sometimes sufficient to convert a rodent cell to a tumorigenic state, many human cancers observed clinically have seven or eight such mutations. The use of genomics to study chemical carcinogenesis is relatively new, but the differences at present appear to be a matter of degree rather than kind.

Differences in regulation of cell division are one possible reason for variation between species in the site of action of a carcinogen, or its potency at a particular site. A finding of carcinogenesis in the mouse liver, for instance, is a reasonably good indicator of potential for carcinogenesis at some site in the human, but not usually in human liver (Huff, 1999). The mouse liver (and to a lesser extent that of the rat) is a common site of spontaneous tumors. It is also relatively sensitive to chemical carcinogenesis. The human liver is apparently more resistant to carcinogenesis; human liver tumors are unusual except when associated with additional predisposing disease, such as hepatitis B or alcoholic cirrhosis. Conversely, other tumor sites are more sensitive in the human. Interspecies variation in site and sensitivity to carcinogenesis may also arise from differences in pharmacokinetics and metabolism, especially for carcinogens where metabolic activation or detoxification is important. This variability may cause important differences in sensitivity between individuals in a diverse population such as humans. Variability between individuals in both susceptibility and pharmacokinetics or metabolism is probably less in experimental animal strains that are bred for genetic homogeneity.

Animal carcinogenesis studies are often designed to maximize the chances of detecting a positive effect, and do not necessarily mimic realistic human exposure scenarios. Thus extrapolation from an experimentally accessible route to that of interest for a risk assessment may be necessary. Even for studies by realistic routes such as oral or inhalation, doses may be large compared to those commonly encountered in the environment, in order to counter the limitation in statistical power caused by the relatively small size of an animal experiment. Whereas the exposed population of an epidemiological study might number in the thousands, a typical animal study might have fifty individuals per exposure group. With this group size any phenomenon with an incidence of less than about 5% is likely to be undetectable. Statistically significant results may be obtained even with groups as small as ten animals per dose group, when incidence of a tumor that is rare in the controls approached 100% in a treated group. The consensus experimental design for animal carcinogenesis studies, which has evolved over the last 50 years of investigation, is represented by the protocol used by the U.S. National Toxicology Program (NTP) for studies using oral routes (diet, gavage or drinking water) or inhalation. These carcinogenesis bioassays usually involve both sexes of an experimental species, and most often two species. NTP has standardized the use of the C57BlxC3H $F_1$ hybrid mouse, and the Fischer 344 rat as the standard test species, although NTP has announced plans to substitute use of the Wistar Han rat for the Fisher 344 rat. There is now an extensive database of background tumor incidences, normal physiology, biochemistry, histology and anatomy for these strains, which aids in the interpretation of pathological changes observed in experiments. Nevertheless, there is enough variation in background rates of common tumors that the use of concurrent controls is essential for hazard identification or dose-response assessment. “Historical control” data are mainly used to reveal anomalous outcomes in the concurrent controls. The fact that a significantly elevated incidence of a tumor relative to the concurrent control group is within the range of historical controls at that site for the test sex and strain is not necessarily grounds for dismissing the biological significance of the finding.

Groups of fifty animals of each sex and species are used, with control groups, and several dose groups, the highest receiving the maximum tolerated dose (MTD). Recent study designs have emphasized the desirability of at least three dose levels covering a decade with "logarithmic" spacing (*i.e.* MTD, 1/2 MTD or 1/3 MTD, and 1/10 MTD). This extended design is aimed at providing better dose-response information, and may contribute important additional information, such as mechanistic insights, for the hazard identification phase.

*Supporting evidence: genetic toxicity, mechanistic studies*

Investigators have developed additional data sources that can support or modify the conclusions of animal carcinogenesis bioassays, and provide information on mechanisms of action of agents suspected of being carcinogenic based on epidemiological studies or animal bioassays.

Genetic damage in exposed organisms includes both gene mutations (point or frameshift), and larger scale effects such as deletions, gene amplification, sister-chromatid exchanges, translocations and loss or duplication of segments or whole chromosomes. These genetic effects of chemical exposures are deleterious in their own right. In addition, since carcinogenesis results from somatic mutations and similar genetic alterations, agents that cause genetic damage generally have carcinogenic potential. Conversely, many known carcinogens are also known to be genotoxic, although there is also a significant class of carcinogens that are not directly genotoxic according to the usual tests. These latter agents presumably work by some other mechanism, although recent genetic studies have shown that even tumors induced by these agents may show mutations, deletions or amplification of growth regulatory genes.

Experimental procedures to demonstrate and measure genetic toxicity may involve exposure of intact animals, and examination of genetic changes in, for example, bone marrow cells (or cells descended from these e.g. the micronucleus test, which detects remnants of chromosomal fragments in immature erythrocytes), mutations in *Drosophila* flies, or appearance of color spots in the coat of mice. However, many tests have employed single celled organisms or mammalian cells in culture. The best known of these tests is the *Salmonella* reverse mutation assay, popularly known as the Ames test after its inventor. This is representative of a larger class of tests for mutagenic activity in prokaryotic organisms (bacteria), which necessarily only look at gene-level mutations. Similar tests in eukaryotic microorganisms (yeasts, *Aspergillus*) and cultured mammalian cells also detect chromosomal effects. Many tests using microorganisms *in vitro* involve addition of activating enzymes (*e.g.* liver postmitochondrial supernatant – "S9") to mimic the metabolism of promutagenic chemicals *in vivo*. Another type of test examines the induction in mammalian cells of morphological transformation or anchorage-independent growth. These changes are seen as analogous to the conversion of a normal tissue cell into a tumorigenic cell.

These various genetic tests contribute different information, which may be used to amplify and confirm conclusions drawn from human studies or animal bioassays, or to draw conclusions in the absence of epidemiological or bioassay data. In the latter case they have also been used in prioritizing agents for further evaluation by means of bioassays.

<u>*Carcinogen Identification schemes*</u>

Some regulatory programs, such as California's Safe Drinking Water and Toxics Enforcement Act ("Proposition 65") and various activities of the U.S. EPA, require that explicit lists of substances having the potential to act as human carcinogens be maintained. Other such lists are developed by non-regulatory research organizations, such as the U.S. National Toxicology Program and the International Agency for Research on Cancer (IARC), an international program of the World Health Organization. The California air toxics programs do not have any statutory requirement to "identify" carcinogens. The requirement instead is to identify hazardous substances as Toxic Air Contaminants, and to determine whether or not a threshold concentration, below which no adverse effects are expected, is likely to exist:

> ***HEALTH AND SAFETY CODE, Division 26 (Air Resources), § 39660.***
>
> *(2) The evaluation shall also contain an estimate of the levels of exposure that may cause or contribute to adverse health effects. If it can be established that a threshold of adverse health effects exists, the estimate shall include both of the following factors:*
>
> *(A) The exposure level below which no adverse health effects are anticipated.*
>
> *(B) An ample margin of safety that accounts for the variable effects that heterogeneous human populations exposed to the substance under evaluation may experience, the uncertainties associated with the applicability of the data to human beings, and the completeness and quality of the information available on potential human exposure to the substance. In cases in which there is no threshold of significant adverse health effects, the office shall determine the range of risk to humans resulting from current or anticipated exposure to the substance.*

In practice however this requirement amounts to the need to establish whether or not a substance is carcinogenic. Any such effect is clearly harmful, and whereas the great majority of non-cancer health effects of chemicals are regarded as having a threshold, the default assumption for carcinogens is that there is no threshold (as described below). OEHHA follows the guidelines laid out by IARC for identification and classification of potential human carcinogens, which are described in detail in the most recent revision of the *Preamble* to the IARC monographs series (IARC, 2006). The IARC Monograph series provides evaluations of the carcinogenicity of individual substances or commonly occurring mixtures. The evaluation guidelines used are similar to those used by other scientific or regulatory authorities, including U.S.EPA.

The data inputs to hazard identification for carcinogens are human epidemiological studies, animal bioassays, along with supporting evidence such as mechanistic and genotoxicity data and structure-activity comparisons. IARC also assembles data on the structure and identity of the agent. The list of agents considered includes specific chemicals and also complex mixtures, occupational and lifestyle factors, physical and biological agents, and other potentially carcinogenic exposures.

IARC evaluations determine the quality of evidence for both animal and human evidence as falling into one of four categories: sufficient evidence of carcinogenicity, limited evidence of carcinogenicity, inadequate evidence of carcinogenicity and evidence suggesting lack of carcinogenicity. Stringent requirements for data quality are imposed. In view of their crucial importance, these definitions are quoted directly from the *Preamble* (IARC 2006):

**"(a) Carcinogenicity in humans**

***Sufficient evidence of carcinogenicity:*** The Working Group considers that a causal relationship has been established between exposure to the agent and human cancer. That is, a positive relationship has been observed between the exposure and cancer in studies in which chance, bias and confounding could be ruled out with reasonable confidence. A statement that there is *sufficient evidence* is followed by a separate sentence that identifies the target organ(s) or tissue(s) where an increased risk of cancer was observed in humans. Identification of a specific target organ or tissue does not preclude the possibility that the agent may cause cancer at other sites.

***Limited evidence of carcinogenicity*****:** A positive association has been observed between exposure to the agent and cancer for which a causal interpretation is considered by the Working Group to be credible, but chance, bias or confounding could not be ruled out with reasonable confidence.

***Inadequate evidence of carcinogenicity:*** The available studies are of insufficient quality, consistency or statistical power to permit a conclusion regarding the presence or absence of a causal association between exposure and cancer, or no data on cancer in humans are available.

***Evidence suggesting lack of carcinogenicity:*** There are several adequate studies covering the full range of levels of exposure that humans are known to encounter, which are mutually consistent in not showing a positive association between exposure to the agent and any studied cancer at any observed level of exposure. The results from these studies alone or combined should have narrow confidence intervals with an upper limit close to the null value (e.g. a relative risk of 1.0). Bias and confounding should be ruled out with reasonable confidence, and the studies should have an adequate length of follow-up. A conclusion of *evidence suggesting lack of carcinogenicity* is inevitably limited to the cancer sites, conditions and levels of exposure, and length of observation covered by the available studies. In addition, the possibility of a very small risk at the levels of exposure studied can never be excluded.

**(b) Carcinogenicity in experimental animals**

Carcinogenicity in experimental animals can be evaluated using conventional bioassays, bioassays that employ genetically modified animals, and other in-vivo bioassays that focus on one or more of the critical stages of carcinogenesis. In the absence of data from conventional long-term bioassays or from assays with neoplasia as the end-point, consistently positive results in several models that address several stages in the multistage process of carcinogenesis should be considered in evaluating the degree of evidence of carcinogenicity in experimental animals.

The evidence relevant to carcinogenicity in experimental animals is classified into one of the following categories:

***Sufficient evidence of carcinogenicity*****:** The Working Group considers that a causal relationship has been established between the agent and an increased incidence of malignant neoplasms or of an appropriate combination of benign and malignant neoplasms in (a) two or more species of animals or (b) two or more independent studies in one species carried out at different times or in different laboratories or under different protocols. An increased incidence of tumours in both sexes of a single species in a well-conducted study, ideally conducted under Good Laboratory Practices, can also provide *sufficient evidence*.

A single study in one species and sex might be considered to provide *sufficient evidence of carcinogenicity* when malignant neoplasms occur to an unusual degree with regard to incidence, site, type of tumour or age at onset, or when there are strong findings of tumours at multiple sites.

***Limited evidence of carcinogenicity*****:** The data suggest a carcinogenic effect but are limited for making a definitive evaluation because, e.g. (a) the evidence of carcinogenicity is restricted to a single experiment; (b) there are unresolved questions regarding the adequacy of the design, conduct or interpretation of the studies; (c) the agent increases the incidence only of benign neoplasms or lesions of uncertain neoplastic potential; or (d) the evidence of carcinogenicity is restricted to studies that demonstrate only promoting activity in a narrow range of tissues or organs.

***Inadequate evidence of carcinogenicity*****:** The studies cannot be interpreted as showing either the presence or absence of a carcinogenic effect because of major qualitative or quantitative limitations, or no data on cancer in experimental animals are available.

***Evidence suggesting lack of carcinogenicity*****:** Adequate studies involving at least two species are available which show that, within the limits of the tests used, the agent is not carcinogenic. A conclusion of *evidence suggesting lack of carcinogenicity* is inevitably limited to the species, tumour sites, age at exposure, and conditions and levels of exposure studied."

IARC utilizes the evaluations of animal and human data, along with supporting evidence including genotoxicity, structure-activity relationships, and identified mechanisms, to reach an overall evaluation of the potential for carcinogenicity in humans. The revised *Preamble* (IARC, 2006) includes a description of the data evaluation criteria for this supporting evidence, and indications as to the situations where the availability of supporting evidence may be used to modify the overall conclusion from that which would be reached on the basis of bioassay and/or epidemiological evidence alone. The overall evaluation is expressed as a numerical grouping, the categories of which are described below, as before by directly quoting IARC (2006):

**"Group 1: The agent is *carcinogenic to humans*.**

This category is used when there is *sufficient evidence of carcinogenicity* in humans. Exceptionally, an agent may be placed in this category when evidence of carcinogenicity in humans is less than *sufficient* but there is *sufficient evidence of carcinogenicity* in experimental animals and strong evidence in exposed humans that the agent acts through a relevant mechanism of carcinogenicity.

**Group 2.**

This category includes agents for which, at one extreme, the degree of evidence of carcinogenicity in humans is almost *sufficient*, as well as those for which, at the other extreme, there are no human data but for which there is evidence of carcinogenicity in experimental animals. Agents are assigned to either Group 2A (*probably carcinogenic to humans*) or Group 2B (*possibly carcinogenic to humans*) on the basis of epidemiological and experimental evidence of carcinogenicity and mechanistic and other relevant data. The terms *probably carcinogenic* and *possibly carcinogenic* have no quantitative significance and are used simply as descriptors of different levels of evidence of human carcinogenicity, with *probably carcinogenic* signifying a higher level of evidence than *possibly carcinogenic*.

**Group 2A: The agent is *probably carcinogenic to humans*.**

This category is used when there is *limited evidence of carcinogenicity* in humans and *sufficient evidence of carcinogenicity* in experimental animals. In some cases, an agent may be classified in this category when there is *inadequate evidence of carcinogenicity* in humans and *sufficient evidence of carcinogenicity* in experimental animals and strong evidence that the carcinogenesis is mediated by a mechanism that also operates in humans. Exceptionally, an agent may be classified in this category solely on the basis of *limited evidence of carcinogenicity* in humans. An agent may be assigned to this category if it clearly belongs, based on mechanistic considerations, to a class of agents for which one or more members have been classified in Group 1 or Group 2A.

**Group 2B: The agent is *possibly carcinogenic to humans*.**

This category is used for agents for which there is *limited evidence of carcinogenicity* in humans and less than *sufficient evidence of carcinogenicity* in experimental animals. It may also be used when there is *inadequate evidence of carcinogenicity* in humans but there is *sufficient evidence of carcinogenicity* in experimental animals. In some instances, an agent for which there is *inadequate evidence of carcinogenicity* in humans and less than *sufficient evidence of carcinogenicity* in experimental animals together with supporting evidence from mechanistic and other relevant data may be placed in this group. An agent may be classified in this category solely on the basis of strong evidence from mechanistic and other relevant data.

**Group 3: The agent is *not classifiable as to its carcinogenicity to humans*.**

This category is used most commonly for agents for which the evidence of carcinogenicity is *inadequate* in humans and *inadequate* or *limited* in experimental animals.

Exceptionally, agents for which the evidence of carcinogenicity is *inadequate* in humans but *sufficient* in experimental animals may be placed in this category when there is strong evidence that the mechanism of carcinogenicity in experimental animals does not operate in humans.

Agents that do not fall into any other group are also placed in this category.

An evaluation in Group 3 is not a determination of non-carcinogenicity or overall safety. It often means that further research is needed, especially when exposures are widespread or the cancer data are consistent with differing interpretations.

**Group 4: The agent is *probably not carcinogenic to humans*.**

This category is used for agents for which there is *evidence suggesting lack of carcinogenicity* in humans and in experimental animals. In some instances, agents for which there is *inadequate evidence of carcinogenicity* in humans but *evidence suggesting lack of carcinogenicity* in experimental animals, consistently and strongly supported by a broad range of mechanistic and other relevant data, may be classified in this group."

The IARC hazard evaluation system provides a detailed and generally accepted scheme to classify the strength of evidence as to the possible human carcinogenicity of chemicals and other agents. This includes careful consideration of mechanistic data and other supporting evidence, the evaluation of which is also important to inform selection of models or defaults used in dose response assessment, as is described below. The extended consideration of supporting evidence is in fact the primary difference between more recent versions of the guidance from IARC, and also by other organizations including U.S. EPA, and the original versions of that guidance. In fact, the basic criteria for hazard identification based on bioassay and epidemiological data have not changed substantially in other respects from earlier guidance documents, including that originally published by California (DHS, 1985). Although as noted earlier the California Air Toxics programs do not categorize identified carcinogens, it has generally been the practice to regard any agent with an IARC overall classification in Group 1 or Group 2 as a known or potential human carcinogen. This implies the selection of various policy-based default options, including absence of a threshold in the dose-response curve, unless specific data are available to indicate otherwise. The same basic identification criteria are used by OEHHA scientific staff to determine the appropriate treatment of agents not evaluated by IARC, or for which newer data or revised interpretations suggest that an earlier IARC determination is no longer appropriate.

U.S. EPA has also proposed a scheme for carcinogen hazard identification and strength of evidence classification in their recently finalized Guidelines for Carcinogen Risk Assessment (U.S. EPA, 2005). These principally differ from the IARC guidance in recommending a more extensive narrative description rather than simply a numerical identifier for the identified level of evidence, and also to some degree in the weight accorded to various types of supporting evidence. However, for most purposes they may be regarded as broadly equivalent to the scheme used by IARC, and OEHHA has chosen to cite the IARC (2006) *Preamble* as representing the most up-to-date and generally accepted guidance on this issue.

**Dose Response Assessment**

The dose-response phase of a cancer risk assessment aims to characterize the relationship between an applied dose of a carcinogen and the risk of tumor appearance in a human. This is usually expressed as a cancer slope factor ["potency" – in units of reciprocal dose - usually $(\mu g/kg\text{-body weight.day})^{-1}$ or "unit risk" – reciprocal air concentration – usually $(\mu g/m^3)^{-1}$] for the lifetime tumor risk associated with lifetime continuous exposure to the carcinogen at low doses. Cancer potency factors may also be referred to as “cancer slope factors”. (As will be described later, additional algorithms may need to be applied to determine risk for specific age groups, or at higher doses where toxicokinetic factors have significant effect.) The basic methodologies recommended in this document are similar to those described by U.S. EPA (2005a) in their Carcinogen Risk Assessment Guidelines. This document therefore refers to U.S. EPA (2005a) for explanation of detailed procedures, and will provide only a brief summary except in cases where OEHHA recommendations are different from or more explicit than those of U.S. EPA.

*Interspecies Extrapolation*

The procedures used to extrapolate low-dose human cancer risk from epidemiological or animal carcinogenicity data are generally health-protective in that they determine an upper confidence bound on the risk experienced by an exposed population. As statistical estimates they cannot be regarded as definite predictions of the risk faced by any one specific individual, who might for a variety of reasons, including individual exposure and susceptibility, experience a risk different from the estimate. The risk assessment procedures used aim to include the majority of variability in the general human population within the confidence bound of the estimate, although the possibility that some individuals might experience either lower or even no risk, or a considerably higher risk, cannot be excluded. Additionally, differences may exist between the characteristics of the general public and those of studied populations. For example, healthy workers, the subject of most epidemiological studies, are often found to have lower rates of morbidity and mortality than the general population (Wen et al., 1983; Monson, 1986; Rothman and Greenland, 1998). Most human data are derived from studies of largely male adult workers and risk estimates cannot take into account specific physiological factors of women, children, and older populations that may affect the potency of a carcinogen, including early age-at-exposure.

Dose-response assessment based on environmental epidemiological studies may involve evaluation of health impacts at exposure levels within the range of those measured in the study population. However, more usually the source data are studies of occupationally exposed humans or of animals, in which case the exposures in the study are likely to be much higher than those of concern for risk assessments relating to community or ambient exposures. Further, even when extrapolation from animal species to humans is not required, the general population to which the URF is applied may differ in characteristics relative to the occupational population studied. It is therefore necessary to extrapolate from the available data to the population and exposure range of concern, which is done by using a dose-response model derived from the source data. The models used fall into three main classes; mechanistically based models, empirical models and (where data are lacking to support a true data-based model) default assumptions. The factors affecting the dose-response relationships for carcinogenesis may also be divided into those relating to absorption, distribution, metabolism and excretion on the one hand (*i.e.* toxicokinetics), and those relating to the underlying dose-response characteristics of carcinogenesis at the tissue or cellular level (*i.e.* toxicodynamics). In this sense the problem of dose response assessment for carcinogens is similar to that for non-cancer toxic effects. The toxicokinetic models used may in fact be similar for both situations, but the toxicodynamic models are generally different.

*Intraspecies Extrapolation and Inter-individual Variability*

In estimating the impact of a particular level of exposure to a carcinogen on a target human population, in addition to extrapolation from a test animal group or human study population to an idealized "average" target human (usually a healthy adult male) it is necessary to consider the range of susceptibility in the target population. In the present case this is typically defined as the general population of the State of California, including of course women (some of whom are pregnant), infants and children, the elderly, the sick, and those with genetic polymorphisms or acquired differences which affect their susceptibility to carcinogens. In general it has been assumed that the upper-bound risk estimates obtained from the standard toxicodynamic models described below are sufficiently health-protective to cover the intrinsic variability of the adult human target population, in spite of the fact that these models do not (except in the few cases where an estimate is based on epidemiological data from a large and unselected study group) explicitly address this type of variability (U.S. EPA, 2005a). However, various analyses (Drew et al., 1983; Barton et al., 2005; Appendix J) have suggested that this assumption is inadequate to cover the expected variability within a human population that includes infants and children. Accordingly both U.S. EPA (2005b) and this document (page 28 *et seq.*) now offer guidance on the use of age-specific adjustment factors to allow for the expected greater sensitivity of infants and children to chemical carcinogenesis.

The ability to accommodate human variability with regard to the toxicokinetic factors affecting susceptibility to carcinogens varies with the level of detail used in the particular assessment. If the generic interspecies extrapolation approach based on body weight is used without any explicit toxicokinetic model then the assumption is made, as in the case of toxicodynamic variability, that the overall health-protective assumptions made are sufficient to cover the toxicokinetic variability. On the other hand if explicit models such as those referenced in the following paragraph are used, this variability may be more explicitly accommodated by using parameter values which are taken as point estimates from measured distributions of population values, or by using Monte Carlo techniques to include those distributions in the model (Bois et al., 1996; OEHHA, 1992; 2001b).

*Toxicokinetic Models*

Considerable literature exists showing the importance of understanding the toxicokinetics of carcinogens in understanding their mechanism of action, sites of impact and dose-response relationships. U.S. EPA (2005) in Section 3.1 refers to the importance of identifying an appropriate dose metric for the dose-response analysis. Early cancer risk assessments typically used applied dose as the dose metric, which is adequate in simple cases provided appropriate correction factors are applied for interspecies extrapolation. However, it is often observed that the uptake, metabolism and elimination of the carcinogenic substance (and/or a procarcinogen and metabolites) is non-linear, especially at the higher doses employed in experimental animal studies (Hoel *et al.*, 1983, Gaylor *et al.*, 1994). Extrapolation to lower doses where such relationships tend to linearity (Hattis, 1990) is aided by the use of toxicokinetic models. These may be relatively simple compartment models, or sophisticated "physiologically based pharmacokinetic (PBPK) models" which to a greater or lesser degree model the actual biochemical and physiological events of toxicokinetic importance. Applications of both types of model may be found in various risk assessment documents prepared for the Toxic Air

Contaminants program (and other OEHHA risk assessments); since the details vary widely according to the nature of the chemical and the availability of appropriate kinetic data these general guidelines will defer to those examples rather than attempt a fuller exposition here. Further analysis of the use of toxicokinetic modeling in extrapolation from animals to humans, and in accounting for interindividual variability among adult humans, infants and children is presented in the Air Toxics Hot Spots *Technical Support Document for the Derivation of Noncancer Reference Exposure Levels* (OEHHA, 2007: Public Review Draft). Although this refers to the use of toxicokinetic modeling in non-cancer risk assessment, the primary considerations are similar for cancer risk assessment.

*<u>Toxicodynamic Models</u>*

An early use of mechanistic analysis to support risk assessment was the development of the Armitage-Doll multistage model of dose-response for carcinogenesis. The multistage model was initially developed on theoretical grounds, and by examination of epidemiological and animal data on time to tumor incidence. Subsequent discovery of the molecular biology of proto-oncogenes has provided a basis for explaining the model in terms of actual biological events and systems (Barrett and Wiseman, 1987). This model was developed by Crump and others into the "linearized multistage model", which has been extensively used for carcinogen risk assessment. It leads to a number of partially verifiable predictions, including linearity of the dose-response relationship at low doses, which is observed for many genotoxic carcinogens. It also predicts the form of the dose-response relationship at higher doses, which generally follow a polynomial form (subject to sampling and background corrections) except where other identifiable factors such as pharmacokinetics intervene.

It has been argued that the simple linearized form of the multistage model has limitations as a description of carcinogenic mechanisms, which detract from its usefulness and generality. Cell proliferation is known to be important in the progression of cancer. It may actually be the primary mechanism of action for a few carcinogens, as opposed to the direct modification of DNA by the carcinogen or a metabolite which is assumed to cause the mutational event at each stage in the original multistage description. A cell proliferation model has been developed (Moolgavkar and Knudson, 1981), which retains the concept of an initiating mutational event (in most cases caused by interaction of the chemical with DNA, although it could also be a spontaneous mutation) as in the original multistage model, but also considers proliferation, death or terminal differentiation of both normal and initiated cells. This model is thought to better describe the biological events in carcinogenesis. However, it has not been used extensively in risk assessment because it requires many parameters that are difficult to define and measure (such as proliferation and death rates for various classes of cell). If these cannot be accurately determined, the model has too many free parameters and is not helpful in defining extrapolated values for risk assessment purposes. This highlights a general problem in using mechanistic models in carcinogen risk assessment, which is that the carcinogenesis data themselves are generally insufficient to define fully the dose response curve shape at low doses or provide much mechanistic information. The analysis is therefore supplemented with policy-based assumptions (such as the expectation of linearity at low doses) and, wherever possible, additional experimental measurements relating to the mechanism of action, in order to make meaningful prediction of risk from environmental exposures to humans.

Because of the difficulties in validating simplified mechanistic models such as the basic multistage model, and the additional difficulty of parameter estimation with more complex mechanistic models, the new U.S. EPA guidelines (U.S. EPA, 2005a) and some recent California risk assessments have chosen instead to use a less overtly mechanistic approach. This approach combines benchmark dose methodology (described below) with an explicit choice of the method for low-dose extrapolation, either assuming low-dose linearity or, for certain carcinogens where data indicate that this is appropriate, a "margin of exposure" or safety/uncertainty factor based approach. This benchmark method is now normally recommended for carcinogen dose response analysis, and the results generally differ little from those derived by the linearized multistage model. Although the linearized multistage method is no longer recommended as the default approach for cancer potency estimation it remains a plausible alternative in many cases, and still has useful applications, such as for time-to-tumor analyses for which benchmark methods are not yet widely available. Additionally, a considerable number of existing cancer potencies in this document were derived by this method. Many of these would not be significantly different if calculated by the benchmark approach, and are unlikely to be replaced soon by newly calculated values. The linearized multistage method will therefore also be briefly described here.

*Benchmark dose methodologies*

The use of benchmark dose methodology has been explored by various investigators [including Gaylor et al. (1998); van Landingham et al. (2001) and Crump (1984, 1995, 2002)] as a tool for dose response extrapolation. This has been recommended in regulatory guidelines for both carcinogenic (U.S. EPA, 2005a) and non-carcinogenic (U.S. EPA, 1995) endpoints. The basic approach is to fit an arbitrary function to the observed incidence data, and to select a "point of departure" (POD) (benchmark dose) *within the range of the observed data.* From this a low dose risk estimate or assumed safe level may be obtained by extrapolation, using an assumed function (usually linear) or by application of uncertainty factors. The critical issue here is that no assumptions are made about the nature of the underlying process in fitting the data. The assumptions about the shape of the dose response curve (linear, threshold, etc.) are explicitly confined to the second step of the estimation process, and are chosen on the basis of policy, mechanistic evidence or other supporting considerations. The benchmark chosen is a point at the low end of the observable dose-response curve. Usually a dose at which the incidence of the tumor is 10% is chosen for animal studies, although lower effect levels may be appropriate for large epidemiological data sets. Because real experimental data include variability in the response of individual subjects, and measurement errors, likelihood methodology is applied in fitting the data. A lower confidence bound (usually 95%) of the effective dose ($LED_{10}$), rather than its maximum likelihood estimate (MLE), is used as the point of departure. This properly reflects the uncertainty in the estimate, taking a cautious interpretation of highly variable or error-prone data. It also reflects the instability of MLE values from complex curve-fitting routines, which has been recognized as a problem also with the linearized multistage model.

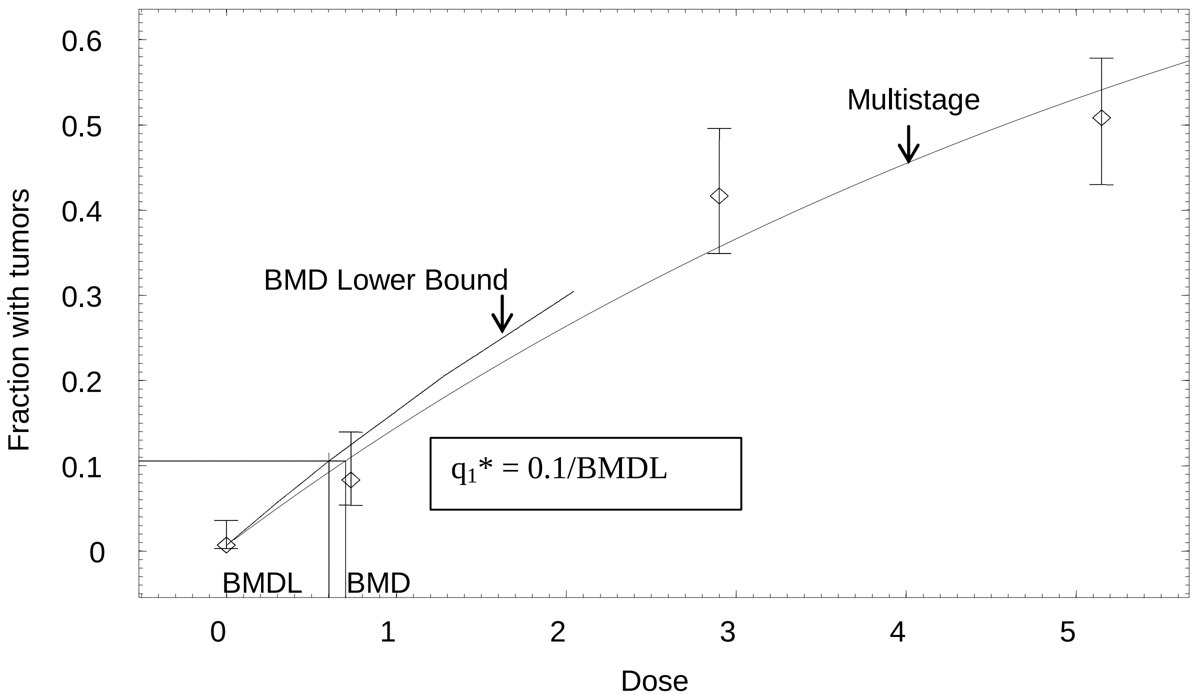
For cancer dose-response estimation using the benchmark dose method, either animal bioassay data or epidemiological data provide a suitable basis. In the absence of a pharmacokinetic model (which could provide tissue-specific dose metrics), the potency would ordinarily be based on the time-weighted average exposure during the exposure or dosing period. The model used to fit the data can be chosen from a range of available alternative quantal models, depending on which provides the best fit to the data in the observable range. In practice, the multistage polynomial fit developed for the linearized multistage model works well for most tumor data sets. Here it is being used merely as a mathematical curve-fitting tool, where the model well fits the data set, without making assumptions about its validity as a biological model of carcinogenesis.

Suitable polynomial fits and estimates of the benchmark may be obtained using U.S. EPA's BMDS software. The benchmark often used is the 95% lower confidence bound on the dose producing 10% tumor incidence. However, if data are available which include a significant dose-response at less than 10% tumor incidence, then that lower benchmark should be used (e.g. $LED_{05}$ or $LED_{01}$). Other software such as Tox_Risk, which was used for the linearized multistage model, has been used successfully, although the earlier GLOBAL program and its relatives are less suitable as curve-fitting tools for benchmark dose analysis.

Since it is usually assumed in cancer risk estimation that the low-dose response relationship is linear, risk estimates and a potency value (slope factor) may be obtained by linear extrapolation from an appropriate benchmark dose. The potency is the slope of that line ($0.1/LED_{10}$). The low dose linearity assumption is a general default for any carcinogen, and it is unlikely to be altered for genotoxic carcinogens.

A calculation using the benchmark dose approach (using a polynomial model with exponents restricted to zero or positive values), and linear extrapolation from the $LED_{10}$ to obtain a potency estimate is shown in Figure 1 (the figure was generated by the U.S. EPA's BMDS program). This is based on tumor incidence data from an actual experiment with vinyl bromide in rats (Benya *et al.*, 1982), with metabolized dose calculated by means of a pharmacokinetic model (Salmon *et al.*, 1992). The value of $q_1$* obtained by this calculation would then be corrected for the duration of the experiment if it had lasted for less than the standard rat lifetime, and for bodyweight and route-specific pharmacokinetic factors as described below. This is in addition to the correction for exposure duration that would be necessary if the study had not lasted for 105 weeks, and the interspecies correction, both of which are described below.

**Figure 1 Benchmark dose calculation for tumor data in rats exposed to vinyl bromide**




From Salmon *et al*. (1992), based on data from Benya *et al*. (1982)

*Linearized Multistage Model*

Quantal analyses

A "multistage" polynomial (U.S. EPA, 1986, 2005a; Anderson *et al*., 1983), based on the mechanistic insights of the original Armitage and Doll model of cancer induction and progression, has been used extensively by U.S. EPA, OEHHA and other risk assessors to model the dose response for lifetime risk of cancer. It usually is used for analysis of animal bioassay data, although related approaches have occasionally been used with epidemiological data. In mathematical terms, the probability of dying with a tumor (P) induced by an average daily dose (d) is:

$$P(d) = 1 - \exp[-(q_0 + q_1 d + q_2 d^2 + ... + q_j d^j)]$$

with constraints

$$q_i \geq 0 \text{ for all i.}$$

Equivalently, $A(d) = 1 - \exp [ - (q_1 d + q_2 d^2 + \cdots + q_k d^k )]$,

where $A(d) = \dfrac{P(d) - P(0)}{1 - P(0)}$ is the extra risk over background at dose *d*.

The $q_i$ model parameters are constants that can be estimated by fitting the polynomial to the data from the bioassay, *i.e.* the number of tumor bearing animals (as a fraction of the total at risk) at each dose level, including the controls. The fit is optimized using likelihood methodology, assuming that the deviations from expected values follow a $\chi^2$ distribution, with the number of degrees of freedom (and hence the maximum number of terms allowed in the polynomial) determined by the number of points in the data set. All the coefficients of the terms are constrained to be zero or positive, so the curve is required to be straight or upward curving, with no maxima, minima or other points of inflection. In addition to the maximum likelihood estimates of the parameters, the upper 95% confidence bounds on these parameters are calculated.

The parameter $q_0$ represents the background lifetime incidence of the tumor. The slope factor $q_1$, or more usually its upper bound ($q_1^*$), is termed the cancer potency. The maximum likelihood estimate (MLE) of $q_1$ is not usually regarded as a reliable estimate for several reasons. First, it fails to reflect the uncertainty and variability in the data which affect the value of the estimate. This is an important issue for protection of public health, which is emphasized by current regulatory guidelines. Secondly, due to the variable order of the polynomial and the effect of some terms being zero as opposed to having a small but finite value, the MLE is unstable, and may show large and unpredictable changes in response to very slight changes in the input data. It may also erratically have a zero value, even when the data imply a significant positive dose-response relationship. The MLE is not a measure of central tendency for this estimate distribution (which is always asymmetrical and often multi-peaked). For small doses, the cancer potency is the ratio of excess lifetime cancer risk to the average daily dose received. Details of the estimation procedure are given in Crump (1981) and Crump, Guess, and Deal (1977). Several software programs are available to perform the necessary calculations, including U.S. EPA's BMDS, Tox_Risk and the earlier GLOBAL programs by Crump and colleagues, and Mstage, written by Crouch (1987).

When dose is expressed in units of mg/kg-d, the potency is given in units of $(\text{mg/kg-d})^{-1}$. As in the case of potencies obtained by the benchmark approach, the experiment-based potency value needs to be corrected for less-than lifetime or intermittent exposure, and extrapolated from the test species to humans. Risk calculations using potency value estimated using the linearized multistage model predict the cancer risk at low doses only, with the higher order terms of the fitted polynomial being ignored since their contribution is negligible at low doses.

*Selection of Site and Tumor Type*

In developing cancer potency estimates from animal data, standard practice has been to use dose-response data for the most sensitive tumor site as the basis of the estimate (CDHS, 1985). Where tumors of more than one histological type (e.g. adenomas and carcinomas) are observed at a single site, the combined incidence, i.e. proportion of animals affected with at least one tumor of any of the relevant types, is used for dose-response assessment. The same rules for combining tumor types are generally applied in determining statistical significance for carcinogen identification (IARC, 2006): tumor types considered to represent different stages of progression following initiation of a common original normal cell type are combined, whereas tumor types having different cellular origins are generally not combined by this procedure. Other considerations that may influence choice of site for dose response estimation include the quality of the data (especially, the statistical impact of a high or variable rate of a particular tumor type and site in control animals), and biological relevance to humans. However, it is an important principle that, just as for the hazard identification phase, concordance of site or tumor type between animal models and human health effects may occur but is not assumed or required.

*Carcinogens inducing tumors at multiple sites*

For most carcinogens, the selection of the most sensitive site in the animal studies is recognized as providing a risk estimate which is appropriate to protect human health. However, for chemicals that induce tumors at multiple sites, the single-site approach may underestimate the true carcinogenic potential. For example, the overall assessment of cancer risk from cigarette smoking (U.S. DHHS, 1982) or ionizing radiation (NRC, 1990) is not based on risk at one site, such as lung cancer. Instead, total cancer risk is estimated from all the sites at which agent-induced tumors are observed (lung, bladder, leukemia, etc), combined.

For carcinogens that induce tumors at multiple sites and/or with different cell types in a particular species and sex, OEHHA derives the animal cancer potency by probabilistically summing the potencies from the different sites and/or cell types. Using the combined potency distribution takes into account the multisite tumorigenicity and provides a basis for estimating the cumulative risk of all treatment-related tumors.

The linear term ($q_1$) of either the multistage model or the multistage-in-dose, Weibull-in-time model is first estimated based on the dose-response data for each of the treatment-related tumor sites. Statistical distributions, rather than point estimates, are generated at each site by tracing the profile likelihood of the linear term ($q_1$) (Zeise et al., 1991). The distributions of $q_1$ for each of the treatment-related sites are then statistically summed using a Monte Carlo approach and assuming independence. The sum is created by adding the linear term for each tumor site, according to its distribution, through random sampling. The upper 95 percent confidence limit on the summed distribution is taken as the multisite animal cancer potency estimate (McDonald et al., 2003, McDonald and Komulainen, 2005).

OEHHA has applied this approach in several recent dose-response analyses, including that for naphthalene presented in Appendix B of this document.

**Early-Lifestage Cancer Potency Adjustments**

In recent years, there have been growing concerns regarding the exposure of children to environmental chemicals, including the possibility that they may be more susceptible than adults to injury caused by those chemicals. The California Legislature passed the Children's Environmental Health Protection Act (Senate Bill 25, Escutia; Chapter 731, Statutes of 1999; "SB 25") to help address these concerns. Under SB25, OEHHA is mandated to consider infants and children specifically, where data permit, in evaluating the health effects of Toxic Air Contaminants (TACs).

The development of cancer is one of the adverse health effects that may occur in children as a result of exposure to environmental chemicals. The document "Prioritization of Toxic Air Contaminants Under the Children's Environmental Health Protection Act" (OEHHA, 2001a) noted that risks of cancer from exposures to carcinogens occurring from conception through puberty can be different than those from exposures occurring in adulthood. Exposure to a carcinogen early in life may result in a greater lifetime risk of cancer for several reasons:

1. Cancer is a multistage process and the occurrence of the first stages in childhood increases the chance that the entire process will be completed, and a cancer produced, within an individual's lifetime.
2. Tissues undergoing rapid growth and development may be especially vulnerable to carcinogenic agents. During periods of increased cell proliferation there is rapid turnover of DNA, and more opportunity for misrepair of damage (e.g., DNA breaks, crosslinks, adducts) or alterations (e.g., altered DNA methylation) to result in permanent changes to the DNA (e.g., mutations) that may ultimately lead to cancer.
3. During early development, a greater proportion of the body's cells are relatively undifferentiated stem cells, and as such represent a large target population of somatic cells capable of passing along permanent changes to the DNA during future cell divisions.
4. There may be greater sensitivity to hormonal carcinogens early in life since the development of many organ systems is under hormonal control (e.g., male and female reproductive systems, thyroid control of CNS development).
5. Other factors that may play a role in increased cancer risk from exposures during critical developmental periods include differences in immunological activity, intestinal absorption, biliary and kidney excretion, blood and fat distribution, and expression of enzyme systems that activate or detoxify carcinogens.

Data in humans and animals for a variety of carcinogens suggest that exposures to such carcinogens early in life may result in a greater lifetime risk of cancer compared to exposures later in life. Two examples of this effect in humans are ionizing radiation and diethylstilbestrol (DES) carcinogenicity.

Ionizing radiation exposure carries an increased risk of cancer when exposures occur early in life compared to adult exposures for a number of tumor types. Children exposed to ionizing radiation (diagnostic X-rays) *in utero* demonstrate a larger excess of leukemia cases than children exposed to ionizing radiation postnatally (NRC, 1990). Exposure to radioisotopes ($^{131}$I,

$^{137}Cs$, $^{134}Cs$, $^{90}Sr$) as a consequence of the 1986 Chernobyl nuclear accident resulted in an elevated thyroid cancer incidence in children but not adults (Moysich, 2002). Age at irradiation in Hodgkin's disease patients treated with radiotherapy strongly influenced the risk of developing breast cancer. The relative risk (RR) of developing breast cancer was 136 for women treated before 15 years of age, 19 for women 15-24 years of age, and 7 for those 24-29 years of age. In women above 30 years of age, the risk was not increased (Hancock *et al*., 1993).

DES was administered to pregnant women in the 1940s-1960s for the purpose of preventing pregnancy loss. In 1970, Herbst and Scully described 7 cases of vaginal adenocarcinoma (6 cases of the clear-cell type) in women aged 15-22 years. This type of cancer is extremely rare in that age range. A follow-up epidemiological study included an additional case, and noted the fact that the mothers of 7 of the 8 patients had been treated with DES during their pregnancy (Herbst *et al*., 1971). Reports by other investigators confirmed the association between maternal use of DES during pregnancy and the development of vaginal adenocarcinoma in their female offspring (Preston-Martin, 1989). It was observed that *in utero* DES exposure resulted in female genital tract morphological changes which correlated with both dose and duration of exposure, and those changes were not related to the maternal conditions which were the reason for the DES administration. Additionally, the risk of occurrence of those morphological changes declined with increasing gestational age at first exposure (O'Brien *et al*., 1979; Preston-Martin, 1989). In contrast, vaginal adenocarcinoma incidence did not increase in the exposed mothers themselves, indicating an increased early-life susceptibility to the carcinogenic effects of DES.

Data from animal studies provide additional examples of increased sensitivity to early life (typically postnatal and juvenile) exposures. These effects span a wide range of target tissues, including the liver (vinyl chloride, safrole), brain (methylnitrosourea), reproductive tract (DES, tamoxifen), and lung (urethane) (OEHHA, 2001a).

U.S. EPA addressed the potential for increased susceptibility to cancer caused by environmental chemicals when the exposure occurs during an early lifestage in "Supplemental Guidance for Assessing Susceptibility from Early-Life Exposure to Carcinogens" (U.S. EPA, 2005b) (referred to henceforth as the Supplemental Guidance). This document is intended to be a companion to the revised "Guidelines for Carcinogen Risk Assessment" (U.S. EPA, 2005a).

U.S. EPA oral exposure cancer risk methodology relies on estimation of the lifetime average daily dose, which can account for exposure factor differences between adults and children (e.g. eating habits and body weight). However, early lifestage susceptibility differences have not been taken into consideration when cancer potency factors were calculated. The Supplemental Guidance document focused on studies that define the potential duration and degree of increased susceptibility that may arise from early-life exposures. An analysis of those studies including a detailed description of the procedures used was published in Barton *et al*. (2005) (included as Appendix I). The criteria used to decide if a study could be included in the quantitative analysis are as follows (excerpted from U.S. EPA, 2005b):

1. Exposure groups at different post-natal ages in the same study or same laboratory, if not concurrent (to control for a large number of potential cross-laboratory experimental variables including pathological examinations),

2. Same strain/species (to eliminate strain-specific responses confounding age-dependent responses),

3. Approximately the same dose within the limits of diets and drinking water intakes that obviously can vary with age (to eliminate dose-dependent responses confounding age-dependent responses),

4. Similar latency period following exposures of different ages (to control for confounding latency period for tumor expression with age-dependent responses), arising from sacrifice at >1 year for all groups exposed at different ages, where early-life exposure can occur up to about 7 weeks. Variations of around 10 to 20% in latency period are acceptable,

5. Postnatal exposure for juvenile rats and mice at ages younger than the standard 6 to 8 week start for bioassays; prenatal (*in utero*) exposures are not part of the current analysis. Studies that have postnatal exposure were included (without adjustment) even if they also involved prenatal exposure,

6. "Adult" rats and mice exposure beginning at approximately 6 to 8 weeks old or older, i.e. comparable to the age at initiation of a standard cancer bioassay (McConnell, 1992). Studies with animals only at young ages do not provide appropriate comparisons to evaluate age-dependency of response (e.g., the many neonatal mouse cancer studies). Studies in other species were used as supporting evidence, because they are relatively rare and the determination of the appropriate comparison ages across species is not simple, and

7. Number of affected animals and total number of animals examined are available or reasonably reconstructed for control, young, and adult groups (i.e., studies reporting only percent response or not including a control group would be excluded unless a reasonable estimate of historical background for the strain was obtainable).

Cancer potencies were estimated from a one-hit model (a restricted form of the Weibull time-to-tumor model), which estimates cumulative incidence for tumor onset. U.S. EPA (2005b) compared the estimated ratio of the cancer potency from early-life exposure to the estimated cancer potency from adult exposure. The general form of the equation for the tumor incidence at a particular dose, [P(dose)] is:

$$P(\text{dose}) = 1-[1-P(0)]\exp(-\text{cancer potency}*\text{dose})$$

where P(0) is the incidence of the tumor in controls. The ratio of juvenile to adult cancer potencies at a single site were calculated by fitting this model to the data for each age group. The model fit depended upon the design of the experiment that generated the data. Studies evaluated by U.S. EPA had two basic design types: experiments in which animals were exposed either as juveniles or as adults (with either a single or multiple dose in each period), and experiments in which exposure began either in the juvenile or in the adult period, but once started, continued through life.

The model equations for the first study type are:

$$P_A = P_0 + (1\text{-}P_0)\,(1\text{-}e^{-m_A\delta_A})$$

$$P_J = P_0 + (1\text{-}P_0)\,(1\text{-}e^{-m_A\,e^{\lambda}\,\delta_A})$$

where $A$ and $J$ refer to the adult and juvenile period, respectively, λ is the natural logarithm of the juvenile:adult cancer potency ratio, $P_0$ is the fraction of control animals with the particular tumor type being modeled, $P_x$ is the fraction of animals exposed in age period $x$ with the tumor, $m_A$ is the cancer potency, and $\delta_x$ is the duration or number of exposures during age period $x$.

The goal of the model is to determine λ, which is the logarithm of the estimated ratio of juvenile to adult cancer potencies. This serves as a measure of potential susceptibility for early-life exposure.

For the second study type, the model equations take into account that exposures that were initiated in the juvenile period continue through the adult period. The model equations for the fraction of animals exposed only as adults with tumors in this design are the same as in the first study type, but the fraction of animals whose first exposure occurred in the juvenile period is:

$$P_J = P_0 + (1\text{-}P_0)\,(1\text{-}e^{-m_A\,e^{\lambda}\,(\delta_J - \delta_A)\text{-}m_A\,\delta_A})$$

$\delta_J$ includes the duration of exposure during the juvenile period and the subsequent adult period.

Parameters in these models were estimated using Bayesian methods and all inferences about the ratios were based on the marginal posterior distribution of λ. A complete description of these procedures (including the potential effect of alternative Bayesian priors that were examined) was published in Barton *et al*. (2005) (Appendix I). This method produced a posterior mean ratio of the early-life to adult cancer potency, which is an estimate of the potential susceptibility of early-life exposure to carcinogens. Ratios of greater or less than one indicate greater or less susceptibility from early-life exposure, respectively.

U.S. EPA reviewed several hundred studies reporting information on 67 chemicals or complex mixtures that are carcinogenic via perinatal exposure. Eighteen chemicals were identified which had animal study designs involving early-life and adult exposures in the same experiment. Of those 18 chemicals, there were overlapping subsets of 11 chemicals involving repeated exposures during early postnatal and adult lifestages and 8 chemicals using acute exposures (usually single doses) at different ages. Those chemicals are listed in Table 1.

**Table 1 Chemicals having animal cancer study data available with early-life and adult exposures in the same experiment.**

| Chemical | Study Type |
|---|---|
| Amitrole | repeat dosing |
| Benzidine | repeat dosing |
| Benzo[a]pyrene (BaP) | acute exposure |
| Dibenzanthracene (DBA) | acute exposure |
| Dichlorodiphenyltrichloroethane (DDT) | lifetime exposure, repeat dosing |
| Dieldrin | lifetime exposure, repeat dosing |
| Diethylnitrosamine (DEN) | acute exposure, lifetime exposure |
| Dimethylbenz[a]anthracene (DMBA) | acute exposure |
| Dimethylnitrosamine (DMN) | acute exposure |
| Diphenylhydantoin, 5,5-(DPH) | lifetime exposure, repeat dosing |
| Ethylnitrosourea (ENU) | acute exposure |
| Ethylene thiourea (ETU) | lifetime exposure, repeat dosing |
| 3-Methylcholanthrene (3-MC) | repeat dosing |
| Methylnitrosourea (NMU) | acute exposure |
| Polybrominated biphenyls (PBBs) | lifetime exposure, repeat dosing |
| Safrole | lifetime exposure, repeat dosing |
| Urethane | acute exposure, lifetime exposure |
| Vinyl chloride (VC) | repeat dosing |

U.S. EPA calculated the difference in susceptibility between early-life and adult exposure as the estimated ratio of cancer potency at specific sites from early-life exposure over the cancer potency from adult exposure for each of the studies that were determined qualitatively to have appropriate study designs and adequate data. The results were grouped into four categories: 1) mutagenic chemicals administered by a chronic dosing regimen to adults and repeated dosing in the early postnatal period (benzidine, diethylnitrosamine, 3-methylcholanthrene, safrole, urethane and vinyl chloride); 2) chemicals without positive mutagenicity data administered by a chronic dosing regimen to adults and repeated dosing in the early postnatal period (amitrole, dichlorodiphenyltrichloroethane (DDT), dieldrin, ethylene thiourea, diphenylhydantoin, polybrominated biphenyls); 3) mutagenic chemicals administered by an acute dosing regimen (benzo[*a*]pyrene, dibenzanthracene, diethylnitrosamine, dimethylbenzanthracene, dimethylnitrosamine, ethylnitrosourea, methylnitrosourea and urethane); 4) chemicals with or without positive mutagenicity data with chronic adult dosing and repeated early postnatal dosing.

The acute dosing animal cancer studies were considered qualitatively useful by U.S. EPA because they involve identical exposures with defined doses and time periods demonstrating that differential tumor incidences arise exclusively from age-dependent susceptibility. However, they were not used to derive a quantitative cancer potency factor age adjustment, primarily because most of the studies used subcutaneous or intraperitoneal injection as a route of exposure. These methods have not been considered quantitatively relevant routes of environmental exposure for human cancer risk assessment by U.S. EPA, for reasons including the fact that these routes of exposure are expected to have a partial or complete absence of first pass metabolism which could affect potency estimates. Additionally, U.S. EPA decided that cancer potency estimates are usually derived from chronic exposures, and therefore, any adjustment to those potencies should be from similar exposures.

The repeated dosing studies with mutagenic chemicals using exposures during early postnatal and adult lifestages were used to develop a quantitative cancer potency factor age adjustment. Studies with repeated early postnatal exposure were included in the analysis even if they also involved earlier maternal and/or prenatal exposure, while studies addressing only prenatal exposure were not used in the analysis. The weighted geometric mean susceptibility ratio (juvenile to adult) for repeated and lifetime exposures in this case was 10.4 (range 0.12 – 111, 42% of ratios greater than 1).

The following adjustments of cancer potency factors were suggested by U.S. EPA, based on the results of the preceding analysis, which concluded that cancer risks generally are higher from early-life exposure than from similar exposure doses and durations later in life:

1. For exposures before 2 years of age (i.e., spanning a 2-year time interval from the first day of birth until a child's second birthday), a 10-fold adjustment.

2. For exposures between 2 and <16 years of age (i.e., spanning a 14-year time interval from a child's second birthday until their sixteenth birthday), a 3-fold adjustment.

3. For exposures after turning 16 years of age, no adjustment.

The 10-fold adjustment used for the 0 – 2 years of age range is approximately the weighted geometric mean cancer potency ratio from juvenile versus adult exposures in the repeated dosing studies. U.S. EPA considered this period to display the greatest toxicokinetic and toxicodynamic differences between children and adults. Data were not available to calculate a specific dose-response adjustment factor for the 2 to <16-year age range, so EPA selected the 3-fold adjustment because it was half the logarithmic scale difference between the 10-fold adjustment for the first two years of life and no adjustment (*i.e.*, 1-fold) for adult exposure. The 3-fold adjustment represents an intermediate level of adjustment applied after 2 years of age through <16 years of age. The upper age limit (16 years of age) reflects the end of puberty and the conclusion of body height growth. U.S. EPA recognizes that the use of a weighted geometric mean of the available study data to develop an age-adjustment factor for cancer potencies may either overestimate or underestimate the actual early-life cancer potency for specific chemicals, and therefore emphasizes in the Supplemental Guidance that chemical-specific data should be used in preference to these default adjustment factors whenever such data are available.

U.S. EPA is recommending the age-dependent adjustment factors described above only for mutagenic carcinogens, because the data for non-mutagenic carcinogens were considered to be too limited and the modes of action too diverse to use this as a category for which a general default adjustment factor approach can be applied. OEHHA considers this approach to be insufficiently health protective. There is no obvious reason to suppose that the toxicokinetics of non-mutagens would be systematically different from those of mutagens. It would also be inappropriate to assume by default that non-mutagenic carcinogens are assumed to need a toxicodynamic correction factor of 1. Most if not all of the factors that make individuals exposed to carcinogens during an early-lifestage potentially more susceptible than those individuals exposed during adulthood also apply to non-mutagenic carcinogen exposures (*e.g.*, rapid growth and development of target tissues, potentially greater sensitivity to hormonal carcinogens, differences in metabolism). It should also be noted that carcinogens that do not cause gene mutations may still be genotoxic by virtue of causing chromosomal damage. Additionally, many carcinogens do not have adequate data available for deciding on a specific mode of action, or do not necessarily have a single mode of action. For these reasons, OEHHA will apply the default cancer potency factor age adjustments described above to all carcinogens unless data are available which allows for the development of chemical-specific cancer potency factor age adjustments. In those cases, an agent-specific model of age dependence (based on observational or experimental data) might be used, or alternative (larger or smaller) adjustment factors and age ranges may be applied where understanding of the mechanism of action and target tissues makes this appropriate.

An independent analysis of animal cancer studies which include early life exposure by the Reproductive and Cancer Hazard Assessment Branch (RCHAB) of OEHHA also supports the application of age-window specific cancer potency factor adjustments. This analysis is provided in Appendix J of this document. Age-related cancer susceptibility data were identified from published animal cancer bioassays in which these issues were addressed. Two types of studies are represented. The first type, "multi-exposure window studies," have exposure groups in at least one of the following age groups - prenatal (from conception to birth), postnatal (from birth to weaning), or juvenile (from weaning to sexual maturity), along with an older age-at-exposure reference group. It should be noted that these age groupings are different than those used by U.S. EPA. For example, prenatal (*in utero*) exposures were not part of the analysis performed by U.S. EPA, and a separate postnatal age group was not included. The second type supports the "case study" of individual chemicals. It includes experiments with at least one group of animals dosed solely during one of the three exposure windows named above. Case studies were constructed from multiple such experiments in each of these early life exposure windows. This document presents results for two case studies: ethyl-N-nitrosoamine (DEN) and N-ethyl-N-nitrosourea (ENU).

Statistical methods were developed and used to analyze the data and derive measures of early-life susceptibility. A cancer potency (the slope of the dose response curve) was developed for each of the experiments selected. When in an experiment, treatment related tumors were observed at multiple sites or at the same site but arising from different cell types, slopes from these sites were statistically combined to create an overall multisite cancer potency. An age sensitivity factor (ASF) is the ratio of cancer potency derived from an early life exposure experiment(s) to that derived from an experiment(s) conducted in adult animals. The ASF can then be used to adjust the existing cancer potency for specific carcinogens. Two ASFs were developed for each experiment. The unadjusted ASF focuses on the inherent susceptibility of the young to the carcinogen and considers potencies for individuals followed for similar periods of time and similarly exposed but for the age window in which the exposure occurs. Therefore, the unadjusted ASF does not address the longer period of time exposed young have to develop tumors. Application of a time-of-dosing adjustment based on the Doll-Armitage model of carcinogenesis is then applied to address this issue. The resulting ASF addresses both the inherent susceptibility to the young to some carcinogens as well as the time-to-tumor issue. Application of this approach for risk associated with lifetime exposures would include an ASF of less than 1 for exposures during the latter part of adult life for carcinogens that act on early stages. Therefore, the addition of this adjustment to the younger lifestages but not to the later part of the adult period could overestimate the risk of whole-life exposures. On the other hand, the 70 year "lifetime" used in estimating lifetime cancer risk does not reflect the longer lifespan of the U.S. population. Further, the animal bioassays on which potency was based typically exclude pre-weaning dosing and sacrifice animals during their late middle-age. Cancer potencies calculated from standard assays can therefore understate lifetime cancer potencies. In reality, the uncertainties in this type of analysis are considerable, but it is important to note that the unadjusted sensitivity ratios underestimate the cancer risk from exposures early in life.

Data from studies on 23 unique carcinogens, 20 of which are considered to act via primarily genotoxic modes of action, were analyzed. Of these 20 carcinogens, 16 are thought to require metabolic activation to the ultimate carcinogenic species (Table 2). The analyses indicate that both the prenatal and postnatal lifestages can be much more susceptible to developing cancer than the adult lifestage. The analyses also indicated that the ASFs for these age windows vary by chemical, gender and species.

For the multi-window studies, the ASF distributions from the underlying studies on the same chemical were combined by equally weighting each chemical, and when multiple ASFs were derived (from multiple studies) on the same chemical, these were sampled by one of three methods, using either weighting by inverse variance (which emphasizes the studies with least variance in their data) or sampling from the ASF distribution with the largest median. The median unadjusted ASF for the postnatal period was 4.6 with either equal weighting or inverse variance weighting and 7.4 when the ASF distribution with the largest median was sampled for each chemical. There were few cases of unadjusted ASFs less than 1 for the postnatal window. Thus, for the chemicals studied, there was greater susceptibility to carcinogens during the early postnatal compared to the adult period. The differences between postnatal and adult susceptibility appear more pronounced once an adjustment is made to the ASF to take into account the longer period cancer has to manifest when exposure occurs early in life. The median adjusted ASF indicated a 13.5-, 13.4- or 21.6- fold greater contribution to lifetime cancer risk (using equal weighting, weighting by inverse variance or selection of the study within a chemical with the largest median, respectively) when exposure occurs during the postnatal period, compared to equivalent exposures occurring during the adult period. The DEN and ENU case studies also exhibited substantial extra susceptibility during the postnatal period.

Regarding exposure *in utero*, few cases were indicative of equal inherent adult and prenatal susceptibility, with an unadjusted ASF of unity. The unadjusted ASF distribution was roughly bimodal, with unadjusted ASFs for several studies significantly greater than unity and several others significantly less than unity. Median (adjusted) ASFs ranged from 2.8 to 7.5, and means ranged from 16.6 to 37.1, depending on the method used to combine studies. This modality in the prenatal ASF distribution was reflected in the DEN and ENU case studies, with results for DEN suggesting inherently less sensitivity than older animals from exposure *in utero*, and for ENU just the opposite. For the DEN and ENU case studies, the referent groups were juvenile rather than adult animals, and the results may have underestimated the ASF, to the extent that some of the apparent sensitivity for DEN and ENU in the early postnatal period carries through to the juvenile period.

**Table 2. Carcinogens for which studies with multi-window exposures are available**

**Genotoxic carcinogens requiring metabolic activation**

Benzidine
Benzo[a]pyrene
Butylnitrosourea
Dibutylnitrosamine
Diethylnitrosamine (DEN)
7,12-Dimethylbenz[a]anthracene (DMBA)
Dimethylnitrosamine (DMN)
Di-n-propylnitrosamine (DPN)
1 -Ethyl-nitrosobiuret
2-Hydroxypropylnitrosamine
3-Hydroxyxanthine
3-Methylcholanthrene (3-MC)
4-(Methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK)
Safrole
Urethane
Vinyl chloride

**Genotoxic carcinogens not requiring metabolic activation**

1,2-Dimethylhydrazine
Ethylnitrosourea (ENU)
Methylnitrosourea (MNU)
ß-Propiolactone

**Nongenotoxic carcinogens**

1,1-Bis(p-chlorophenol)-2,2,2-trichloroethane (DDT)
Diethylstilbestrol (DES)
2,3,7,8-Tetrachlorodibenzodioxin (TCDD)

ENU is a direct acting carcinogen that does not require metabolic activation, whereas DEN can not be metabolized to any significant extent by fetal tissues until relatively late in gestation. This may explain the lower fetal susceptibility of DEN. However, *in utero* metabolic status is not the sole determinant of *in utero* susceptibility; e.g., benzidine and safrole require metabolic activation and exhibit greater susceptibility from *in utero* exposure.

There were only five chemicals and seven studies, two of which were not independent, available to examine susceptibility in the juvenile period. The unadjusted ASF indicated significantly greater susceptibility in this period for independent studies, with the remaining studies consistent with equal inherent susceptibility to adult animals. Median estimates of juvenile adjusted ASFs from the multi-window analysis range from 4.5 to 5.5 and mean estimates from 7.1 to 9.4, depending upon the method used to combine studies.

The studies that comprise the set of multi-window studies available for these analyses were not homogeneous. That is, they do not represent observations from the same distribution. Sensitivity analyses were conducted to test the robustness of the findings to different procedures for analyzing data and combining results. Of the methods used to combine the ASF distributions for underlying studies within each exposure window, the method of equally weighting studies within a chemical appeared to best represent the available data. The use of inverse variance in weighting ASF distributions within a chemical appeared to underweight small studies and overweight large ones. The method of selecting a single study (i.e., that with the largest median ASF) to represent each chemical may also result in a bias toward higher ASF values to the extent that a selected study was not representative.

In adjusting the ASF to take into account the longer period of time for early carcinogen exposures to result in tumors, the hazard function was assumed to increase with the third power of age. If the true rate of increase with age is greater than that, then the use of these ASFs may result in underestimates of the true sensitivity of these early life stages.

As the multi-window and case studies show, there appears to be considerable variability in age-at-exposure related susceptibility across carcinogens. There is also variability in age-at-exposure related susceptibility among studies of the same carcinogen. The sources of variability evident in the analyzed studies include timing of exposure within a given age window, and gender, strain, and species differences in tumor response. The set of studies identified and analyzed was not sufficiently robust to fully describe the variability quantitatively. This variability raises concerns that selection of the median (the $50^{th}$ percentile) estimates may considerably underestimate effects for certain agents or population groups. Relatively large variability in humans in response to carcinogens is expected to be common (Finkel, 1995).

Several of the carcinogens studied induced tumors at multiple sites in the same experiment, and at different sites, depending upon the age window during which exposure occurred. For these cases the combined multisite potency distribution referred to above was the basis for the life stage comparison. This approach differs from other researchers investigating early vs. late in life differences who focused on tumor site-specific measures of carcinogenic activity (e.g., Barton *et al.*, 2005; Hattis *et al.*, 2004, 2005). OEHHA believes that use of combined multisite potency distributions provides a more complete approach for considering age specific differences in carcinogenic activity. However, the observation that early life is generally a period of increased susceptibility was similarly found using the tumor site-specific approach by these other researchers.

One limitation of the approach was the focus on age windows, without attempting to describe changes in susceptibility within an age window. Timing of carcinogen exposure within a given age window can affect the cancer outcome. For example, experiments with 1-ethyl-1-nitroso-biuret in prenatal and adult rats showed a three fold difference in activity between groups exposed on prenatal day -10 versus prenatal day -3. In a second example, female rats exposed early in the adult period were more than three times as sensitive to the breast cancer effects of MNU as females exposed six weeks later. In general, the adult comparison groups in the multi-window studies were fairly young. The extent to which this may result in an overall bias of the results presented here is unclear. Also for several cases, juvenile animals were used as the later life exposure group. In these cases the ASFs are likely underestimates of the relative sensitivity of the prenatal and postnatal lifestages, compared to that of the adult lifestage.

Excluded from the analysis were early in life studies in which exposure of a given exposure group crossed multiple age windows. An example of results from studies of this type is provided by mouse studies for two non-genotoxic carcinogens, diphenylhydantoin (Chhabra *et al*., 1993a) and polybrominated biphenyls (PBBs) (Chhabra *et al*., 1993b), in which exposures began prior to conception, and continued throughout the prenatal, postnatal, and post-weaning period, up to the age of eight weeks. The data demonstrate an increased sensitivity of the early life period. Some studies that crossed multiple age windows were included in the analyses of Barton *et al*. (2005) (Appendix I), which are consistent with the general conclusions discussed above.

OEHHA recognizes the limitations in the data and analyses presented, as discussed above. However, the analyses do provide some guidance on the extent risk may be over or underestimated by current approaches. While there is a great deal of variability across chemicals in the prenatal ASFs, the data indicate that the potency associated with prenatal carcinogen exposure is not zero. A factor of 10 falls roughly at the 70th percentile for the multi-window study analysis. This value could be applied to the potency estimate when calculating lifetime cancer risk in humans arising from carcinogen exposures that occur *in utero*. In view of the considerable variability in the data for different carcinogens and the limited database available for analysis, OEHHA is not including the application of this factor to cancer potency estimates for *in utero* exposures as a default position in these Guidelines. However, the applicability of a cancer potency adjustment factor for *in utero* exposure will be evaluated on a case-by-case basis, and may be used as evidence develops that supports such use. The consideration of prenatal exposures, including application of an appropriate susceptibility factor, would not make a large difference for risk estimates based on continuous lifetime exposures, due to the relatively short duration of gestation. However, risk estimates for short-term or intermittent exposures might be significantly increased by inclusion of the risks to the fetus *in utero*. Thus, risk may be underestimated when this window is excluded from the analysis.

Selection of appropriate values for the ASFs for postnatal and juvenile exposures is complicated by the limited database of chemicals and studies available for analysis, and the broad, multimodal distribution of results for different chemicals as is shown in Appendix J. In view of the variability thus shown, and the considerable uncertainty in applying conclusions from this relatively small set of chemicals to the much larger number of chemicals of concern, it is probably unreasonable to specify a default ASF with greater than half-log precision (*i.e.* values of 1, 3, 10, 30 etc.). A factor of 10 for postnatal exposures falls just below the median estimate for postnatal studies. This is also the value selected by U.S. EPA; while it is consistent with the

OEHHA analysis, it may underestimate risk for some chemicals. The multimodal nature of the distribution of observed chemical-specific sensitivity ratios clearly indicates certain number of chemicals for which the sensitivity ratio is much larger than 10: further research is needed to develop criteria for identifying these cases. Similarly, a factor of 3 for juvenile exposures is consistent with the range of estimates derived from the multi-window studies, although it falls below the median estimates for all three methods presented. It is acknowledged that there are few data available on which to base an estimate for the juvenile period. A factor of 3 adjusts for the longer time available for cancer to manifest, but may not fully account for some inherent differences in susceptibility to cancer, for example those observed in breast tissue of pubescent girls exposed to radiation. For specific carcinogens where data indicate enhanced sensitivity during age windows other than the immediate postnatal and juvenile periods, or demonstrate sensitivity ratios different from the default ASFs, the chemical-specific data should be used in order to adequately protect public health.

## Other Source Documents for Cancer Risk Assessment Guidance

As noted previously, the cancer potencies and unit risks tabulated in this technical support document have been developed by various programs over a number of years. The methods used therefore necessarily varied according to the date of the assessment and the program responsible. The following section summarizes the sources and procedures most commonly applied, and their historical context where this is apposite.

*United States Environmental Protection Agency (U.S. EPA)*

The U.S. EPA was one of the first regulatory agencies to develop and apply cancer risk assessment methodology. Their guidance documents and technical publications have been influential for many programs, including the California Air Toxics (Toxic Air Contaminants and Hot Spots) programs.

*Guidelines for Carcinogen Risk Assessment (U.S. EPA, 1986)*

Prior to the more recent guidelines updating project which, after nearly ten years of internal and public review drafts culminated in the 2005 final revision (see below), U.S. EPA carcinogen risk assessment procedures were generally as described in Anderson *et al*. (1983) and “Guidelines for Carcinogen Risk Assessment” (U.S. EPA, 1986). These methods, which are outlined below, were used to calculate the Integrated Risk Information System (IRIS) cancer potency values, some of which are cited in this document. U.S. EPA has always indicated that cancer risk estimates based on adequate human epidemiologic data are preferred if available over estimates based on animal data. Although the newer guidelines offer alternative methods for dose-response analysis of animal bioassays, and updated consideration of specific topics such as lifestage-related differences in sensitivity, and mechanism of action for some types of carcinogen, the underlying principles, and many of the specific procedures developed in these original guidelines are still applicable and in use.

U.S. EPA Calculation of Carcinogenic Potency Based on Animal Data

In extrapolating low-dose human cancer risk from animal carcinogenicity data, it is generally assumed that most agents that cause cancer also damage DNA, and that the quantal type of biological response characteristic of mutagenesis is associated with a linear non-threshold dose-response relationship. U.S. EPA stated that the risk assessments made with this model should be regarded as conservative, representing the most plausible upper limit for the risk. The mathematical expression used by U.S. EPA in the 1986 guidelines to describe the linear non-threshold dose-response relationship at low doses is the linearized multistage procedure developed by Crump (1980). This model is capable of fitting almost any monotonically increasing dose-response data, and incorporates a procedure for estimating the largest possible linear slope at low extrapolated doses that is consistent with the data at all experimental dose levels. A description of the linearized multistage procedure has been provided above (page 27). U.S. EPA used an updated version (GLOBAL86, Howe *et al*., 1986) of the computer program GLOBAL79 developed by Crump and Watson (1979) to calculate the point estimate and the 95% upper confidence limit of the extra risk $A(d)$.

U.S. EPA separated tumor incidence data according to organ sites or tumor types. The incidence of benign and malignant tumors was combined whenever scientifically defensible. U.S. EPA considered this incidence combination scientifically defensible unless the benign tumors are not considered to have the potential to progress to the associated malignancies of the same histogenic origin. The primary comparison in carcinogenicity evaluation is tumor response in dosed animals as compared to contemporary matched control animals. However, U.S. EPA stated that historical control data could be used along with concurrent control data in the evaluation of carcinogenic responses, and notes that for the evaluation of rare tumors, even small tumor responses may be significant compared to historical data. If several data sets (dose and tumor incidence) are available (different animal species, strains, sexes, exposure levels, exposure routes) for a particular chemical, the data set used in the model was the set where the incidence is statistically significantly higher than the control for at least one test dose level and/or where the tumor incidence rate shows a statistically significant trend with respect to dose level. The data set generating the highest lifetime cancer risk estimate ($q_1^*$) was chosen where appropriate. An example of an inappropriate data set would be a set which generates an artifactually high risk estimate because of a very small number of animals used. If there are 2 or more data sets of comparable size for a particular chemical that are identical with respect to species, strain, sex and tumor sites, the geometric mean of $q_1^*$ estimated from each of those data sets was used for risk estimation. U.S. EPA assumed that mg/surface area/day is an equivalent dose between species. Surface area was further assumed to be proportional to the 2/3 power of the weight of the animal in question. Equivalent dose was therefore computed using the following relationship:

$$\mathrm{d} = \frac{1_e * m}{L_e * W^{2/3}}$$

where $L_e$ = experimental duration, $l_e$ = exposure duration, $m$ = average dose (mg/day) and $W$ = average animal weight. Default average body weights for humans, rats and mice are 70, 0.35 and 0.03 kg, respectively.

Exposure data expressed as ppm in the diet were generally converted to mg/day using the relationship $m$ = ppm * $F$ * $r$, where ppm is parts per million of the chemical in the diet, $F$ is the weight of the food consumed per day in kg, and $r$ is the absorption fraction (assumed to be 1 in the absence of data indicating otherwise). The weight of food consumed, calories required, and animal surface area were generally all considered to be proportional to the 2/3 power of the animal weight, so:

$$m \propto \text{ppm} * W^{2/3} * r, \text{ or } \quad \frac{m}{rW^{2/3}} \propto \text{ppm}$$

The relationship could lead to the assumption that dietary ppm is an equivalent exposure between species. However, U.S. EPA did not believe that this assumption is justified, since the calories/kg food consumed by humans is significantly different from that consumed by laboratory animals (primarily due to differences in moisture content). An empirically derived food factor, $f = F/W$ was used, which is the fraction of a species' body weight consumed per day as food. U.S. EPA (1986) gave the $f$ values for humans, rats and mice as 0.028, 0.05 and 0.13, respectively. Dietary exposures expressed as concentrations in ppm were converted to mg/surface area using the following relationship:

$$\frac{m}{r * W^{2/3}} = \frac{\text{ppm} * F}{W^{2/3}} = \frac{\text{ppm} * f * W}{W^{2/3}} = \text{ppm} * f * W^{2/3}$$

Exposures expressed as mg/kg/day ($m/Wr = s$) were converted to mg/surface area using the relationship:

$$\frac{m}{rW^{2/3}} = s * W^{2/3}$$

The calculation of dose when exposure is via inhalation was performed for cases where 1) the chemical is either a completely water-soluble gas or aerosol and is absorbed proportionally to the amount of inspired air, or 2) where the chemical is a partly water-soluble gas which reaches an equilibrium between the inspired air and body compartments. After equilibrium is attained, the rate of absorption is proportional to metabolic rate, which is proportional to the rate of oxygen consumption, which is related to surface area.

Exposure expressed as mg/day to completely water-soluble gas or aerosols can be calculated using the expression $m = I * v * r$, where $I$ is the inspiration rate/day in $m^3$, $v$ is the concentration of the chemical in air (mg/$m^3$), and $r$ is the absorption fraction (assumed to be the same for all species in the absence of data to the contrary; usually 1). For humans, the default inspiration rate of 20 $m^3$ has been adopted. Inspiration rates for 113 g rats and 25 g mice have been reported to be 105 and 34.5 liters/day, respectively. Surface area proportionality can be used to determine inspiration rate for rats and mice of other weights; for mice, $I = 0.0345\ (W / 0.025)^{2/3}$ $m^3$/day; for rats, $I = 0.105\ (W / 0.113)^{2/3}$ $m^3$/day. The empirical factors for air intake/kg/day (i) for humans, rats and mice are 0.29, 0.64 and 1.3, respectively. Equivalent exposures in mg/surface area can be calculated using the relationship:

$$\frac{m}{W^{2/3}} = \frac{Ivr}{W^{2/3}} = \frac{iWvr}{W^{2/3}} = iW^{1/3}vr$$

Exposure expressed as mg/day to partly water-soluble gases is proportional to surface area and to the solubility of the gas in body fluids (expressed as an absorption coefficient $r$ for that gas). Equivalent exposures in mg/surface area can be calculated using the relationships $m = kW^{2/3} * v * r$, and $d = m/W^{2/3} = kvr$. The further assumption is made that in the case of route-to-route extrapolations (e.g., where animal exposure is via the oral route, and human exposure is via inhalation, or vice versa), unless pharmacokinetic data to the contrary exist, absorption is equal by either exposure route.

Adjustments were made for experimental exposure durations shorter than the lifetime of the test animal; the slope $q_1^*$ was increased by the factor $(L/L_e)^3$, where $L$ is the normal lifespan of the experimental animal and $L_e$ is the duration of the experiment. This assumed that if the average dose $d$ is continued, the age-specific rate of cancer will continue to increase as a constant function of the background rate. Since age-specific rates for humans increase by at least the 2nd power of the age, and often by a considerably higher power (Doll, 1971), there is an expectation that the cumulative tumor rate, and therefore $q_1^*$, will increase by at least the 3rd power of age. If the slope $q_1^*$ is calculated at age $L_e$, it would be expected that if the experiment was continued for the full lifespan $L$ at the same average dose, the slope $q_1^*$ would have been increased by at least $(L/L_e)^3$.

U.S. EPA Calculation of Carcinogenic Potency Based on Human Data

U.S. EPA stated that existing human epidemiologic studies with sufficiently valid exposure characterization are always used in evaluating the cancer potency of a chemical. If they showed a carcinogenic effect, the data were analyzed to provide an estimate of the linear dependence of cancer rates on lifetime cancer dose (equivalent to the factor $q_1^*$). If no carcinogenic effect was demonstrated and carcinogenicity had been demonstrated in animals, then it was assumed that a risk does exist, but it is smaller than could have been observed in the epidemiologic study. An upper limit of cancer incidence was calculated assuming that the true incidence is just below the level of detection in the cohort studied, which is largely determined by the cohort size. Whenever possible, human data are used in preference to animal data. In human epidemiologic studies, the response is measured as the relative risk of the exposed cohort of individuals compared to the control group. The excess risk ($R(X)$ - 1, where $R(X)$ is relative risk) was assumed to be proportional to the lifetime average exposure $X$, and to be the same for all ages. The carcinogenic potency is then equal to $[R(X) - 1]/X$ multiplied by the lifetime risk at that site in the general population. According to this original procedure, the confidence limit for the excess risk was not usually calculated: this decision was ascribed to the difficulty in accounting for inherent uncertainty in the exposure and cancer response data. More recent assessments have taken the opposite view and attempted to calculate and characterize this uncertainty by determining confidence limits, *inter alia*.

*Guidelines for Carcinogen Risk Assessment (U.S. EPA, 2005a)*

U.S. EPA revised its “Guidelines for Carcinogen Risk Assessment” (referred to henceforth as the “U.S. EPA Guidelines”) in 2005. Compared to the 1986 version of this document, more emphasis is placed on establishing a "mode of action" (MOA). The following excerpt provides a definition of this term:

> "The term "mode of action" is defined as a sequence of key events and processes, starting with interaction of an agent with a cell, proceeding through operational and anatomical changes, and resulting in cancer formation. A "key event" is an empirically observable precursor step that is itself a necessary element of the mode of action or is a biologically based marker for such an element. Mode of action is contrasted with "mechanism of action," which implies a more detailed understanding and description of events, often at the molecular level, than is meant by mode of action".

Cancer risk assessments performed under the prior U.S. EPA Guidelines sometimes included a MOA description. However, the 1986 U.S. EPA Guidelines did not explicitly mandate the development of a MOA description in cancer risk assessments.

The MOA information is then used to govern how a cancer risk assessment shall proceed. Tumor incidence data sets arising from a MOA judged to be not relevant to humans are not used to extrapolate a cancer potency factor. If an MOA cannot be determined or is determined to have a low-dose linear dose-response and a nonmutagenic MOA, then a linear extrapolation method is used to develop a cancer potency factor. The same linear extrapolation is used for all lifestages, unless chemical specific information on lifestage or population sensitivity is available. Carcinogens that act via an MOA judged to have a nonlinear low-dose dose response are modeled using MOA data, or the RfD/RfC risk assessment method is used as a default. Adjustments for susceptible lifestages or populations are to be performed as part of the risk assessment process.

If a carcinogen is deemed to act via a mutagenic MOA, then the data from the MOA analysis is evaluated to determine if chemical-specific differences between adults and juveniles exist and can be used to develop a chemical-specific risk estimate incorporating lifestage susceptibility. If this cannot be done, then early-life susceptibility is assumed, and age-dependent adjustment factors (ADAFs) (described above, page 28 et seq.) are applied as appropriate to develop risk estimates. In cases where it is not possible to develop a toxicokinetic model to perform cross-species scaling of animal tumor data sets which arise from oral exposures, the U.S. EPA Guidelines state that administered doses should be scaled from animals to humans on the basis of equivalence of $mg/kg^{3/4}$-d (milligrams of the agent normalized by the 3/4 power of body weight per day). This is a departure from the 1986 U.S. EPA guidelines, which used a 2/3 power of body weight normalization factor. Other adjustments for dose timing, duration and route are generally assumed to be handled in similar fashion to that described for the 1986 guidelines, although of course updated parameter values would be used where available.

The 2005 U.S. EPA Guidelines also use benchmark dose methodology (described above, page 25) to develop a "point-of departure" (POD) from tumor incidence data. For linear extrapolation, the POD is used to calculate a cancer potency factor, and for nonlinear extrapolation the POD is used in the calculation of a reference dose (RfD) or reference concentration (RfC).

It should be noted that none of the cancer potency factors listed in this document were obtained from U.S. EPA risk assessments performed under the 2005 U.S. EPA Guidelines. All U.S. EPA IRIS cancer potency values contained in this document were obtained from risk assessments using the 1986 U.S. EPA Guidelines.

*Office of Environmental Health Hazard Assessment (OEHHA),California Environmental Protection Agency*

The cancer risk assessment procedures originally used by the Office of Environmental Health Hazard Assessment (OEHHA) are outlined in "Guidelines for Chemical Carcinogen Risk Assessments and their Scientific Rationale" (referred to below as the Guidelines) (CDHS, 1985). These procedures were generally used in generating Toxic Air Contaminant (TAC) cancer potency values, standard Proposition 65 cancer potency values and Public Health Goal (PHG) cancer potency values. Expedited Proposition 65 cancer potency values depart somewhat from those procedures and are discussed separately below.

OEHHA cancer risk assessment methodology as described by CDHS (1985) generally resembled that used at that time by U.S. EPA (Anderson *et al*., 1983; U.S. EPA, 1986). OEHHA risk assessment practice similarly reflects the evolution of the technical methodology (e.g. as described in U.S. EPA, 2005a) since the original guidelines were published. The basic principles and procedures described below are still considered applicable: more recent additions to OEHHA cancer risk assessment methods such as the use of benchmark dose methodologies and early-lifestage cancer potency adjustments are discussed above. The Guidelines state that both animal and human data, when available, should be part of the dose-response assessment.

*OEHHA Calculation of Carcinogenic Potency Based on Animal Data*

The procedures used to extrapolate low-dose human cancer risk from animal carcinogenicity data assumed that a carcinogenic change induced in a cell is transmitted to successive generations of cells descended from that cell, and that the initial change in the cell is an alteration (e.g. mutation, rearrangement, etc.) in the cellular DNA. Non-threshold models are used to extrapolate to low-dose human cancer risk from animal carcinogenicity data.

Several models were proposed for extrapolating low-dose human cancer risk from animal carcinogenicity data in the original Guidelines; these models include the Mantel-Bryan method (log-probit model), the one-hit model, the linearized multistage procedure, the gamma multihit model, and a number of time-to-tumor models. The Guidelines stated that time-to-tumor models (i.e., a Weibull-in-time model) should be used for low-dose extrapolation in all cases where supporting data are available, particularly when survival is poor due to competing toxicity. However, the Guidelines also noted the difficulty of determining the actual response times in an experiment; internal tumors are generally difficult to detect in live animals and their presence is usually detected only at necropsy. Additionally, use of these models often requires making the determination of whether a tumor was the cause of death, or was found only coincidentally at necropsy when death was due to other causes. Further, competing causes of death, such as chemical toxicity, may decrease the observed time-to-tumor for nonlethal cancers by allowing earlier necropsy of animals in higher dose groups. The linearized multistage (LMS) procedure was noted as being an appropriate method for dose extrapolation in most cases, with the primary exception being a situation in which sufficient empirical data are available to indicate a dose-response curve of a “quasi-threshold” type (e.g., flat for two or three dose levels, then curving sharply upwards). In this case, the LMS procedure may underestimate the number of stages and overestimate the low-dose risks. In this case, the gamma multihit model was suggested as being a potential alternative. The Mantel-Bryan model was described as having little biological basis as applied to carcinogenesis, and being likely to underestimate risks at low doses. The Guidelines stated that this model should not be used for low dose extrapolation. More recent practice has departed from these original guidelines in some respects, for instance by experimenting with cell-proliferation based models in a few cases: however the LMS model remained the preferred extrapolation model for most purposes. Some of the difficulties in achieving a satisfactory fit to tumor incidence data were found to be alleviated by application of toxicokinetic models and use of an internal rather than applied dose metric with the LMS model. This has resulted in the alternative models originally advocated (Gamma multihit, Mantel-Bryan) being mostly abandoned. As noted above (Dose-Response Assessment, page 21), the use of allegedly biologically based statistical models such as LMS has fallen from favor in recent years, and benchmark dose methodology has become the preferred method for extrapolating cancer potency values from animal cancer incidence data. However, it should also be noted that results generated by the LMS model and benchmark dose methodology from the same data set are often quite similar.

The 1985 Guidelines stated that both animal and human data, when available, should be part of the dose-response assessment. Although preference was given to human data when these were of adequate quality, animal studies may provide important supporting evidence. Low-dose extrapolation of human cancer risk from animal carcinogenicity data was generally based on the most sensitive site, species and study demonstrating carcinogenicity of a particular chemical, unless other evidence indicates that the data set in question is not appropriate for use. Where both benign and malignant tumors are induced at the same site and the malignant tumors are significantly increased, the data on both types of tumors could be combined to form the basis for risk assessment. Pharmacokinetic data on chemical metabolism, effective dose at target site, or species differences between laboratory test animals and humans were considered in dose-response assessments when available. In performing exposure scaling from animals to humans, the “surface area” correction (correcting by the 2/3 power of body weight) was used unless specific data indicates that this should not be done. The Guidelines assumed that in the absence of evidence to the contrary, chemicals that cause cancer after exposure by ingestion will also cause cancer after exposure by inhalation, and vice versa. These original proposals have continued in use with little change except that currently, TAC and PHG cancer potency factor calculations use a 3/4 power of body weight correction for interspecies scaling, in line with current U.S. EPA practice. The standard Proposition 65 cancer potency factor calculations still use a 2/3 power correction because the cancer potency calculation method is specified in regulation (California Health and Safety Code 25249.5 *et seq*.).

Cancer unit risk factors [in units of $(\mu g/m^3)^{-1}$ ]have been calculated from cancer potency factors [in units of $(mg/kg\text{-}day)^{-1}$ ] using the following relationship:

$$UR = \frac{CPF * 20\ m^3}{70\ kg\ *\ CV}$$

where UR is the cancer unit risk, CPF is the cancer potency factor, 70 kg is the reference human body weight, 20 $m^3$ is the reference human inspiration rate/day, and CV is the conversion factor from mg to µg (= 1000). The cancer unit risk describes the excess cancer risk associated with an inhalation exposure to a concentration of 1 µg/$m^3$ of a given chemical; the cancer potency factor describes the excess cancer risk associated with exposure to 1 mg of a given chemical per kilogram of body weight.

It should be noted that although this default method is still used in deriving published cancer unit risk values, for site-specific risk assessments age-appropriate distributions and percentile values are used in the current version of the Hot Spots exposure assessment document. Where exposure to children occurs (as it does in most exposures to the general population surrounding a source site) it is also necessary to apply the age-specific adjustment factors for the appropriate durations in accordance with the guidance offered above (Page 28 *et seq.*).

*OEHHA Calculation of Carcinogenic Potency Based on Human Data*

Human epidemiologic studies with adequate exposure characterization are used to evaluate the cancer potency of a chemical. If they show a carcinogenic effect, the data are analyzed to provide an estimate of the linear dependence of cancer rates on lifetime cancer dose. The 1985 Guidelines stated that with continuous exposure, age-specific incidence continues to increase as a power function (e.g., $t^3$ or $t^4$) of the elapsed time since initial exposure. Lifetime risks can be estimated by applying such a power function to the observed data and extrapolating beyond the actual followup period. OEHHA has generally undertaken the calculation of study power and confidence bounds on the potency estimate as important tools to establish the credibility of the estimate obtained and in comparing this with other estimates (from other human studies or from animal data). Due to the diversity in quality and type of epidemiological data, the specific approaches used in OEHHA risk assessments based on human epidemiologic studies vary on a case by case basis rather than following explicit general guidelines. Examples of the methods used can be observed in the Toxic Air Contaminant documents (these documents are listed in Appendix D: the methods used are described in the compound summaries provided in Appendix B).

*Expedited Proposition 65 Cancer Risk Assessment Methodology*

Expedited cancer potency values developed for several agents listed as carcinogens under Proposition 65 (California Health and Safety Code 25249.5 *et seq.*) were derived from selected animal carcinogenicity data sets of the Carcinogenic Potency Database (CPDB) of Gold *et al.* (1984, 1986, 1987, 1989, 1990) using default procedures specified in the administrative regulations for Proposition 65 (Title 22 California Code of Regulations [CCR] 12703). OEHHA hazard assessments usually describe all relevant data on the carcinogenicity (including dose-response characteristics) of the chemical under examination, followed by an evaluation of any pharmacokinetic and mechanistic (e.g. genotoxicity) data. An evaluation of the data set for the chemical may indicate that adjustments in target dose estimates or use of a dose response model different from the default are appropriate. The procedure used to derive expedited Proposition 65 cancer potency values differs from the usual methodology in two ways. First, it relies on cancer dose response data evaluated and extracted from the original literature by Gold *et al.* Second, the choice of a linearized multistage procedure for generating cancer potency values is automatic, and pharmacokinetic adjustments are not performed. The methods used to develop expedited cancer potency values incorporate the following assumptions:

1. The dose response relationship for carcinogenic effects in the most sensitive species tested is representative of that in humans.

2. Observed experimental results can be extrapolated across species by use of the interspecies factor based on "surface area scaling."

3. The dose to the tissue giving rise to a tumor is assumed to be proportional to the administered dose.

4. The linearized multistage polynomial procedure can be used to extrapolate potency outside the range of experimental observations to yield estimates of "low" dose potency.

5. Cancer hazard increases with the third power of age.

The Carcinogenic Potency Database of Gold *et al.* (1984, 1986, 1987, 1989, 1990) contains the results of more than 4000 chronic laboratory animal experiments on 1050 chemicals by combining published literature with the results of Federal chemical testing programs (Technical Reports from the Carcinogenesis Bioassay Program of the National Cancer Institute (NCI)/National Toxicology Program (NTP) published prior to June 1987). The published literature was searched (Gold *et al*., 1984) through the period December 1986 for carcinogenicity bioassays; the search included the Public Health Service publication "Survey of Compounds Which Have Been Tested for Carcinogenic Activity" (1948-1973 and 1978), monographs on chemical carcinogens prepared by the International Agency for Research on Cancer (IARC) and Current Contents. Also searched were Carcinogenesis Abstracts and the following journals: British Journal of Cancer, Cancer Letters, Cancer Research, Carcinogenesis, Chemosphere, Environmental Health Perspectives, European Journal of Cancer, Food and Cosmetics Toxicology, Gann, International Journal of Cancer, Journal of Cancer Research and Clinical Oncology (formerly Zeitschrift fur Krebsforschung und Klinische Onkologie), Journal of Environmental Pathology and Toxicology, Journal of Toxicology and Environmental Health, Journal of the National Cancer Institute, and Toxicology and Applied Pharmacology. Studies were included in the database if they met the following conditions:

1. The test animals were mammals.
2. Chemical exposure was started early in life (100 days of age or less for hamsters, mice and rats).
3. Route of administration was via the diet, drinking water, gavage, inhalation, intravenous injection or intraperitoneal injection.
4. The test chemical was administered alone (not in combination with other chemicals).
5. Chemical exposure was chronic, with not more than 7 days between exposures.
6. The experiment duration was at least half the standard lifespan for the species used.
7. The study design included a control group and at least 5 animals/exposure group.
8. No surgical interventions were performed.
9. Pathology data were reported for the number of animals with tumors (not total number of tumors).
10. All results reported were original data (not analysis of data reported by other authors).

Included in their data set tabulations are estimates of average doses used in the bioassay, resulting tumor incidences for each of the dose levels employed for sites where significant responses were observed, dosing period, length of study and histopathology. Average daily dose levels were calculated assuming 100% absorption. Dose calculations follow procedures similar to those of Cal/EPA and U.S. EPA; details on methods used and standard values for animal lifespans, body weights, and diet, water and air intake are listed in Gold *et al.* (1984). OEHHA (1992) reviewed the quality assurance, literature review, and control procedures used in compiling the data and found them to be sufficient for use in an expedited procedure. Cancer potency estimates were derived by applying the mathematical approach described in the section below to dose response data in the Gold *et al.* database. The following criteria were used for data selection:

1. Data sets with statistically significant increases in cancer incidence with dose ($p \leq 0.05$) were used. (If the authors of the bioassay report considered a statistically significant result to be unrelated to the exposure to the carcinogen, the associated data set was not used.)
2. Data sets were not selected if the endpoint was specified as "all tumor-bearing animals" or results were from a combination of unrelated tissues and tumors.
3. When several studies were available, and one study stood out as being of higher quality due to numbers of dose groups, magnitude of the dose applied, duration of study, or other factors,

the higher quality study was chosen as the basis for potency calculation if study results were consistent with those of the other bioassays listed.

4. When there were multiple studies of similar quality in the sensitive species, the geometric mean of potencies derived from these studies was taken. If the same experimentalists tested two sexes of the same species/strain under the same laboratory conditions, and no other adequate studies were available for that species, the data set for the more sensitive sex was selected.

5. Potency was derived from data sets that tabulate malignant tumors, combined malignant and benign tumors, or tumors that would have likely progressed to malignancy.

Cancer potency was defined as the slope of the dose response curve at low doses. Following the default approach, this slope was estimated from the dose response data collected at high doses and assumed to hold at very low doses. The Crump linearized multistage polynomial (Crump *et al.*, 1977) was fit to animal bioassay data:

$$\text{Probability of cancer} = 1 - \exp[-(q_0 + q_1 d + q_2 d^2 + ...)]$$

Cancer potency was estimated from the upper 95 % confidence bound on the linear coefficient $q_1$, which is termed $q_1^*$.

For a given chemical, the model was fit to a number of data sets. As discussed in the section above, the default was to select the data for the most sensitive target organ in the most sensitive species and sex, unless data indicated that this was inappropriate. Deviations from this default occur, for example, when there are several bioassays or large differences exist between potency values calculated from available data sets.

Carcinogenicity bioassays using mice and/or rats will often use an exposure duration of approximately two years. For standard risk assessments, this is the assumed lifespan for these species. Animals in experiments of shorter duration are at a lower risk of developing tumors than those in the standard bioassay; thus potency is underestimated unless an adjustment for experimental duration is made. In estimating potency, short duration of an experiment was taken into account by multiplying $q_1^*$ by a correction factor equal to the cube of the ratio of the assumed standard lifespan of the animal to the duration of the experiment ($T_e$). This assumes that the cancer hazard would have increased with the third power of the age of the animals had they lived longer:

$$q_{animal} = q_1^* \ * (104 \text{ weeks}/T_e)^3$$

In some cases excess mortality may occur during a bioassay, and the number of initial animals subject to late occurring tumors may be significantly reduced. In such situations, the above described procedure can, at times, significantly underestimate potency. A time-dependent model fit to individual animal data (i.e., the data set with the tumor status and time of death for each animal under study) may provide better potency estimates. When Gold *et al.* indicated that survival was poor for a selected data set, a time-dependent analysis was attempted if the required data were available in the Tox Risk (Crump *et al.*, 1991) data base. The Weibull multistage model (Weibull-in-time; multistage-in-dose) was fit to the individual animal data.

To estimate human cancer potency, $q_{animal}$ values derived from bioassay data were multiplied by an interspecies scaling factor (K; the ratio of human body weight ($bw_h$) to test animal body weight ($bw_a$), taken to the 1/3 power (Anderson *et al.*, 1983)):

$$K = (bw_h/bw_a)^{1/3}$$

Thus, cancer potency = $q_{human} = K * q_{animal}$

## Chemical-specific Descriptions of Cancer Potency Value Derivations

Unit Risk and potency values for chemicals whose cancer potency values were obtained from Toxic Air Contaminant documents, standard or expedited Proposition 65 documents, U.S. EPA's Integrated Risk Information System (IRIS) documents and Health Effects Assessment Summary Table (HEAST) entries, or from other documents prepared by OEHHA's Air Toxicology and Epidemiology Branch or Pesticide and Environmental Toxicology Branch are presented in Appendix A. Information summaries for these chemicals are presented in Appendix B.

## REFERENCES

Anderson EL and the Carcinogen Assessment Group of the U.S. Environmental Protection Agency. 1983. Quantitative approaches in use to assess cancer risk. Risk Anal 3:277-295.

Barrett JC and Wiseman RW. 1987. Cellular and molecular mechanisms of multistep carcinogenesis: relevance to carcinogen risk assessment. Environ Health Perspect 76:65-70.

Barton HA, Cogliano VJ, Flowers L, Valcovic L, Setzer RW and Woodruff TJ. 2005. Assessing susceptibility from early-life exposure to carcinogens. Environ Health Perspect 113:1125-1133.

Benya TJ, Busey WM, Dorato MA and Berteau PE. 1982. Inhalation carcinogenicity bioassay of vinyl bromide in rats. Toxicol Appl Pharmacol 64:367-379.

Bois FY, Gelman A, Jiang J, Maszle DR, Zeise L and Alexeeff G. 1996. Population toxicokinetics of tetrachloroethylene. Arch Toxicol 70:347-55.

California Department of Health Services (CDHS). 1985. Guidelines for Chemical Carcinogen Risk Assessments and Their Scientific Rationale. CDHS, Health and Welfare Agency, Sacramento, CA.

California Environmental Protection Agency (Cal/EPA). 1992. Expedited Cancer Potency Values and Proposed Regulatory Levels for Certain Proposition 65 Carcinogens. Office of Environmental Health Hazard Assessment, Reproductive and Cancer Hazard Assessment Section, Berkeley, CA.

Crouch E. 1992. MSTAGE (Version 1.1). E.A.C. Crouch, Cambridge Environmental Inc., 58 Buena Vista Road, Arlington, Massachusetts 02141.

Crouch E and Wilson R. 1979. Interspecies comparison of carcinogenic potency. J Toxicol Environ Health 5:1095-1118.

Crump KS. 1980. An improved procedure for low-dose carcinogenic risk assessment from animal data. J Environ Pathol Toxicol 5:675-684.

Crump KS. 1984. A new method for determining allowable daily intakes. Fundam Appl Toxicol 4:854-871.

Crump KS. 1995. Calculation of benchmark doses from continuous data. Risk Anal 15:78-89.

Crump KS. 2002. Critical issues in benchmark calculations from continuous data. Crit Rev Toxicol 32:133-153.

Crump KS, Guess HA and Deal LL. 1987. Confidence intervals and test of hypotheses concerning dose response relations inferred from animal carcinogenicity data. Biometrics 33:437-451.

Crump KS, Howe RB, Van Landingham C and Fuller WG. 1991. TOXRISK Version 3. TOXicology RISK Assessment Program. KS Crump Division, Clement International Division, 1201 Gaines Street, Ruston LA 71270.

Crump KS and Watson WW. 1979. GLOBAL79: A FORTRAN program to extrapolate dichotomous animal carcinogenicity data to low doses. National Institute of Environmental Health Sciences, Contract No. 1-ES-2123.

Doll R. 1971. Weibull distribution of cancer: implications for models of carcinogenesis. J Royal Stat Soc A 13:133-166.

Drew RT, Boorman GA, Haseman JK, McConnell EE, Busey WM and Moore JA. 1983. The effect of age and exposure duration on cancer induction by a known carcinogen in rats, mice, and hamsters. Toxicol Appl Pharmacol 68:120-130.

Finkel AM. 1995. Toward less misleading comparisons of uncertain risks: the example of aflatoxin and alar. Environ Health Perspect 103:376-385

Freireich EJ, Gehan EA, Rall DP, Schmidt LH and Skipper HE. 1966. Quantitative comparison of toxicity of anticancer agents in mouse, rat, hamster, dog, monkey, and man. Cancer Chemother Rep 50:219-244.

Gaylor DW, Gold LS. 1994. Quick estimate of the regulatory virtually safe dose based on the maximum tolerated dose for rodent bioassays. Regul Toxicol Pharmacol 22:57-63.

Gaylor D, Ryan L, Krewski D and Zhu Y. 1998. Procedures for calculating benchmark doses for health risk assessment. Regul Toxicol Pharmacol 28:150-164.

Gold L, de Veciana M, Backman G, Magaw R, Lopipero P, Smith M, Blumenthal M, Levinson R, Bernstein L and Ames B. 1986. Chronological supplement to the Carcinogenic Potency Database: Standardized results of animal bioassays published through December 1984 and by the National Toxicology Program through May 1986. Environ Health Perspect 74:237-329.

Gold L, Sawyer C, Magaw R, Backman G, de Veciana M, Levinson R, Hooper N, Havender W, Bernstein L, Peto R, Pike M and Ames B. 1984. A Carcinogenic Potency Database of the standardized results of animal bioassays. Environ Health Perspect 58:9-319.

Gold L, Slone T, Backman G, Eisenberg S, Da Costa M, Wong M, Manley N and Ames B. 1990. Third chronological supplement to the Carcinogenic Potency Database; Standardized results of animal bioassays published through December 1986 and by the National Toxicology Program through June 1987. Environ Health Perspect 84:215-285.

Gold L, Slone T, Backman G, Magaw R, Da Costa M and Ames B. 1987. Second chronological supplement to the Carcinogenic Potency Database; Standardized results of animal bioassays published through December 1984 and by the National Toxicology Program through May 1986. Environ Health Perspect 74:237-329.

Gold L, Slone T, and Bernstein L. 1989. Summary of carcinogenic potency and positivity for 492 rodent carcinogens in the Carcinogenic Potency Database. Environ Health Perspect 79:259-272.

Hancock SL, Tucker MA and Hoppe RT. 1993. Breast cancer after treatment of Hodgkin's disease. J Natl Cancer Inst 85:25-31.

Hattis D. 1990. Pharmacokinetic principles for dose-rate extrapolation of carcinogenic risk from genetically active agents. Risk Anal 10:303-16.

Hattis D, Goble R, Russ A, Chu M, Ericson J. 2004. Age-related differences in susceptibility to carcinogenesis: a quantitative analysis of empirical animal bioassay data. Environ Health Perspect 112:1152-1158.

Hattis D, Goble R, Chu M. 2005. Age-related differences in susceptibility to carcinogenesis. II. Approaches for application and uncerta1inty analyses for individual genetically acting carcinogens. Environ Health Perspect 113:509-16.

Herbst AL and Scully RE. 1970. Adenocarcinoma of the vagina in adolescence. A report of 7 cases including 6 clear-cell carcinomas (so-called mesonephromas). Cancer 25:745-757.

Herbst AL, Ulfelder H and Poskanzer DC. 1971. Adenocarcinoma of the vagina. Association of maternal stilbestrol therapy with tumor appearance in young women. N Engl J Med 284:878-881.

Hill AB. 1971. Statistical evidence and inference. In: Principles of Medical Statistics, 9th ed., pp. 309-323. Oxford University Press, New York, NY.

Hoel DG, Kaplan NL, Anderson MW. 1983. Implication of nonlinear kinetics on risk estimation in carcinogenesis. Science 219:1032-1037.

Howe RB, Crump KS and Van Landingham C. 1986. GLOBAL86: A computer program to extrapolate quantal animal toxicity data to low doses. Clement Associates, Inc., Ruston, LA.

Huff J. 1999. Long-term chemical carcinogenesis bioassays predict human cancer hazards. Issues, controversies, and uncertainties. Ann N Y Acad Sci 895:56-79.

IARC. 2006. Monographs on the Evaluation of Carcinogenic Risks to Humans: Preamble. International Agency for Research on Cancer, Lyon, France. Available at: http://monographs.iarc.fr/ENG/Preamble/CurrentPreamble.pdf

Institute of Medicine (2004). Gulf War and Health: Updated literature review of Sarin. The National Academy of Sciences, National Academy Press, Washington, DC, pp 20-22. www.nap.edu

Lilienfeld AM and Lilienfeld DE. 1980. Foundations of epidemiology. Oxford University Press, Oxford, England.

McConnell EE. 1992. Comparative response in carcinogenesis bioassay as a function of age at first exposure. In: Guzelian P, Henry CJ, Olin SS, eds. Similarities and difference between children and adults: implications for risk assessment. ILSI Press, Washington, DC, pp. 66−67.

McDonald T, Hoover S, Faust J, Rabovsky J, MacGregor MK, Sherman C, Sandy M and Zeise L. 2003. Development of cancer potency estimates for California's Proposition 65. Poster at Society of Toxicology Annual Meeting. March 2003, Salt Lake City, UT. Abstract No. 687, Toxicological Sciences 72, S-1, 142.

McDonald T and Komulainen H. 2005. Carcinogenicity of the chlorination disinfection by-product MX. Journal of Environmental Science and Health Part C, 23:163–214.

Monson RR. 1986. Observations on the healthy worker effect. J Occup Med 28: 425-433.

Moolgavkar SH, Knudson AG Jr. 1981. Mutation and cancer: a model for human carcinogenesis. J Natl Cancer Inst 66:1037-1052.

Moysich KB, Menezes RJ and Michalek AM. 2002. Chernobyl-related ionising radiation exposure and cancer risk: an epidemiological review. Lancet Oncol 3:269-279.

National Research Council (NRC). 1983. Risk Assessment in the Federal Government: Managing the Process. Committee on the Institutional Means for Assessment of Risks to Public Health. National Academy Press, Washington, DC.

National Research Council (NRC). 1990. Health Effects of exposure to low levels of ionizing radiation. BEIR V. Committee on the Biological Effects of Ionizing Radiation. National Academy Press, Washington, DC

National Research Council (NRC). 1994. Science and Judgment in Risk Assessment. Committee on Risk Assessment of Hazardous Air Pollutants, Board on Environmental Studies and Toxicology, Commission on Life Sciences. National Academy Press, Washington, DC.

O'Brien PC, Noller KL, Robboy SJ, Barnes AB, Kaufman RH, Tilley BC and Townsend DE. 1979. Vaginal epithelial changes in young women enrolled in the National Cooperative Diethylstilbestrol Adenosis (DESAD) project. Obstet Gynecol 53:300-308.

Office of Environmental Health Hazard Assessment (OEHHA). 1992. Proposed Identification of Perchloroethylene as a Toxic Air Contaminant. Part B: Health Effects. (Approved by the Scientific Review Panel, 1991: revised 1992). California Environmental Protection Agency, Sacramento, CA.

Office of Environmental Health Hazard Assessment (OEHHA). 2001a. Prioritization of Toxic Air Contaminants Under the Children's Environmental Health Protection Act. California Environmental Protection Agency, Sacramento, CA.

Office of Environmental Health Hazard Assessment (OEHHA). 2001b. Public Health Goals for chemicals in drinking water: Tetrachloroethylene. California Environmental Protection Agency, Sacramento, CA.

Office of Environmental Health Hazard Assessment (OEHHA). 2005. Air Toxics Hot Spots Program Risk Assessment Guidelines. Part II: Technical Support Document for Describing Available Cancer Potency Factors. California Environmental Protection Agency, Sacramento, CA.

Office of Environmental Health Hazard Assessment (OEHHA). 2005. Proposed Identification of Environmental Tobacco Smoke as a Toxic Air Contaminant. Part B: Health Effects. As approved by the Scientific Review Panel, June 24, 2005. California Environmental Protection Agency, Sacramento, CA.

Office of Environmental Health Hazard Assessment (OEHHA). 2007. Air Toxics Hot Spots Technical Support Document for the Derivation of Noncancer Reference Exposure Levels (Public Review Draft). California Environmental Protection Agency, Sacramento, CA.

Preston-Martin S. 1989. Epidemiological studies of perinatal carcinogenesis. IARC Sci Publ 96:289-314.

Rothman K and Greenland S. 1998. Modern epidemiology. 2nd edition. Lippincott–Raven, Philadelphia, PA, pp. 133-134.

Salmon AG, Monserrat L and Brown JP. 1992. Use of a pharmacokinetic model in cancer risk assessment for vinyl bromide. Presented at the Society of Toxicology Annual Meeting, Seattle, WA, February 1992. Abstract: The Toxicologist 12(1): 96.

Travis CC and White RK. 1988. Interspecific scaling of toxicity data. Risk Anal 8:119-125.

U.S. Dept. of Health and Human Services (U.S. DHHS) 1982. The health consequences of smoking: Cancer. A Report of the Surgeon General. United States Department of Health and Human Services. Pub No (PHS) 82-50179. Washington DC.

U.S. Dept. of Health and Human Services (U.S. DHHS). 1994. The health consequences of smoking: a report of the Surgeon General. Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health, Washington, DC.

U.S. Environmental Protection Agency (U.S. EPA). 1986. Guidelines for Carcinogen Risk Assessment. Federal Register 51:33992-34003.

U.S. Environmental Protection Agency. 2002. A review of the reference dose and reference concentration process. Risk Assessment Forum, Washington, DC. EPA/630/P-02/002F. Available from: http://cfpub.epa.gov/ncea/raf/recordisplay.cfm?deid=55365.

U.S. Environmental Protection Agency. 2005a. Guidelines for Carcinogen Risk Assessment. Risk Assessment Forum, Washington, DC. EPA/630/P-03/001F.

U.S. Environmental Protection Agency. 2005b. Supplemental Guidance for Assessing Susceptibility from Early-Life Exposure to Carcinogens. EPA/630/R-03/003F. Available from: http://www.epa.gov/iris/children032505.pdf.

Van Landingham CB, Allen BC, Shipp AM and Crump KS. 2001. Comparison of the EU T25 single point estimate method with benchmark dose response modeling for estimating potency of carcinogens. Risk Anal. 21:641-56.

Wen CP, Tsai SP and Gibson RL. 1983. Anatomy of the healthy worker effect: A critical review. J Occup Med 25: 283-289.

Zeise L, Salmon AG, McDonald T and Painter P. 1991. Cancer potency estimation. In: Risks of carcinogenesis from urethane exposure. Salmon AG and Zeise L, eds, CRC Press, Boca Raton, FL, pp 97-112.